

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:39:51 / Search time 235.2 Seconds
(without alignments)
3623.632 Million cell updates/sec

Title: US-09-889-325-4
Sequence: 1 MERLIDVERLQAMERAFRR.....HLSPHALVGLAVTELLQVAR 1208

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: UniProt 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------------|--------------------|
| 1 | 6424 | 100.0 | 1208 | 1 RECO4 HUMAN | 094761 homo sapien |
| 2 | 3944 | 61.4 | 1216 | 1 RECO4 MOUNR | 075nr7 mus musculi |
| 3 | 2859.5 | 41.4 | 1500 | 2 040YX6 XENLA | 04jnr8 xenopus lae |
| 4 | 2310.5 | 36.0 | 910 | 2 04RLC3 TETNG | 04rlr3 tetraodon n |
| 5 | 1723.5 | 26.8 | 1579 | 2 09VSE6 DROME | 09vse6 drosophila |
| 6 | 1680.5 | 26.2 | 1530 | 2 09NH11 DROME | 09nh11 drosophila |
| 7 | 1670.5 | 26.0 | 1058 | 2 07PMM8 ANOCA | 07pmm8 anopheles g |
| 8 | 1172 | 18.2 | 340 | 2 07YR85 BOVIN | 07yr85 bos taurus |
| 9 | 971.5 | 15.1 | 927 | 2 07X829 ORYSA | 07x829 oryza sativ |
| 10 | 897.5 | 14.0 | 870 | 2 08W028 ARATH | 08w028 arabidopsis |
| 11 | 872 | 13.6 | 941 | 2 02C6N0 ARATH | 02c6n0 arabidopsis |
| 12 | 761 | 11.8 | 874 | 2 07FAH0 ORYSA | 07fah0 oryza sativ |
| 13 | 635 | 9.9 | 704 | 2 07MID7 VIBRY | 07mid7 vibrio vuln |
| 14 | 631.5 | 9.8 | 608 | 1 RECO ECOLI | 08x8t1 escherichia |
| 15 | 631.5 | 9.8 | 611 | 2 08X8T1 ECOS7 | 08x8t1 escherichia |
| 16 | 631.5 | 9.8 | 611 | 2 08F8M6 ECOS6 | 08f8m6 escherichia |
| 17 | 631.5 | 9.8 | 611 | 2 0831W3 SHIFL | 0831w3 shigella fl |
| 18 | 626 | 9.7 | 639 | 2 08DB16 VIBRV | 08db16 vibrio vuln |
| 19 | 617.5 | 9.6 | 654 | 2 087MB5 VIBPA | 087mb5 vibrio para |
| 20 | 613.5 | 9.6 | 608 | 2 06CZM9 ERWCT | 06czm9 erwinia car |
| 21 | 613 | 9.5 | 642 | 2 087J96 PSEEM | 087j96 pseudomonas |
| 22 | 609.5 | 9.5 | 609 | 2 08ZB32 SALTI | 08zb32 salmonella |
| 23 | 609.5 | 9.5 | 615 | 2 057HE9 SALCH | 057he9 salmonella |
| 24 | 607.5 | 9.5 | 608 | 1 RECO SALTY | 08k1z8 pseudomonas |
| 25 | 604 | 9.4 | 644 | 2 04KJ28 PSEFS | 04kj28 pseudomonas |
| 26 | 604 | 9.4 | 649 | 2 04ZL17 PSESY | 04zl17 pseudomonas |
| 27 | 604 | 9.4 | 654 | 2 088EKL SHRON | 088ekl shewanella |
| 28 | 603.5 | 9.4 | 609 | 2 05PRK6 SALPA | 05prk6 salmonella |
| 29 | 599 | 9.3 | 746 | 2 07NH48 GLOVI | 07nh48 gloeobacter |
| 30 | 594.5 | 9.3 | 595 | 2 04NM16 JDDEL | 04nm16 anaeromyob |
| 31 | 593 | 9.2 | 610 | 2 066FY3 YERSIN | 066fy3 yersinia ps |

ALIGNMENTS

| RESULT 1 | RECO4 HUMAN | STANDARD | PRT | 1208 AA. |
|-------------------------------|---|-----------------------------------|--------|----------|
| ID | 094761 | 096DM2 | 096P55 | |
| AC | 30-MAY-2000 | (Rel. 39, Created) | | |
| DT | 30-MAY-2000 | (Rel. 39, Last sequence update) | | |
| DT | 13-SEP-2005 | (Rel. 48, Last annotation update) | | |
| DE | ATP-dependent DNA helicase Q4 (EC 3.6.1.1) (RecQ protein-like 4) | | | |
| GN | Name=RECO4; Synonym=RECQ4; | | | |
| OS | Homo sapiens (Human) | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; | | | |
| OC | Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RC | NUCLEOTIDE SEQUENCE [mRNA], AND TISSUE SPECIFICITY. | | | |
| RC | TISSUE=Testis; | | | |
| RX | MEDLINE=99097344; PubMed=9878247; DOI=10.1006/geno.1998.5595; | | | |
| RA | Kitao S., Ohnogi I., Ichikawa K., Goto M., Furuchi Y., Shimamoto A., | | | |
| RT | "Cloning of two new human helicase genes of the RecQ family: | | | |
| RT | biological significance of multiple species in higher eukaryotes." | | | |
| RL | Genomics 54:443-452 (1998). | | | |
| RN | [2] | | | |
| RC | NUCLEOTIDE SEQUENCE [GENOMIC DNA], SUBCELLULAR LOCATION, AND DISEASE. | | | |
| RX | MEDLINE=20021764; PubMed=10552928; DOI=10.1006/geno.1999.5959; | | | |
| RA | Kitao S., Lindor N.M., Shitatori M., Furuchi Y., Shimamoto A., | | | |
| RT | "Rothmund-Thomson syndrome responsible gene, RECQL4: genomic structure | | | |
| RT | and products." | | | |
| RL | Genomics 61:268-276 (1999). | | | |
| RN | [3] | | | |
| RC | NUCLEOTIDE SEQUENCE [LARGE SCALE mRNA] OF 468-1208. | | | |
| RC | TISSUE=Lymph, and Placenta; | | | |
| RX | MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899; | | | |
| RA | Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G., | | | |
| RT | Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., | | | |
| RT | Altschuld S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K., | | | |
| RA | Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Heich F., | | | |
| RA | Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., | | | |
| RA | Scaplestein M., Soares M.B., Bonaudo M.F., Casavant T.L., Scheetz T.E., | | | |
| RA | Rah S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J., | | | |
| RA | Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., | | | |
| RA | Villarino S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., | | | |
| RA | Villarino D.K., Murray D.M., Sodergren R.J., Lu X., Gibbs R.A., | | | |
| RA | Faney J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A., | | | |
| RA | Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., | | | |
| RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., | | | |
| RA | Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., | | | |
| RA | Butterfield Y.S.N., Krzywiński M.I., Skalka U., Small D.B., | | | |
| RA | Schneerch A., Schein J.E., Jones S.J.M., Marra M.A., | | | |
| RT | "Generation and initial analysis of more than 15,000 full-length human | | | |
| RT | and mouse cDNA sequences." | | | |
| Proc. Natl. Acad. Sci. U.S.A. | 99:16899-16903 (2002). | | | |

[4]
RN INDUCTION.
RX PubMed=11032027; DOI=10.1038/sj.onc.1203841;
RA Kawabe T., Tenuyama N., Kitao S., Nishikawa K., Shimamoto A.,
RA Shizatori M., Matsumoto T., Anno K., Sato T., Mitsu Y., Seki M.,
RA Enomoto T., Goto M., Ellis N.A., Ide T., Furutachi Y., Sugimoto M.,
RT "differential regulation of human RecQ family helicases in cell
transformation and cell cycle.";
RL Oncogene 19:4764-4772(2000).
RN [5]
RX PubMed=12952869; DOI=10.1093/hmg/ddg306;
RA Sittonen H.A., Kopra O., Kaeserlaenen H., Haravuori H., Winter R.M.,
RA Saemmen A.-M., Peltonen L., Kesilae M.,
RT "Molecular defect of RAPADILINO syndrome expands the phenotype
spectrum of RECQL diseases.";
RL Hum. Mol. Genet. 12:2837-2844(2003).
RN [6]
RX INTERACTIONS WITH UBR1 AND UBR2, SUBCELLULAR LOCATION, FUNCTION, AND
RT IDENTIFICATION BY MASS SPECTROMETRY.
RX PubMed=15317757; DOI=10.1093/hmg/ddh269;
RA Yin J., Kwon Y.T., Varshavsky A., Wang W.,
RT "RECQL4, mutated in the Rothmund-Thomson and RAPADILINO syndromes,
interacts with ubiquitin ligases UBR1 and UBR2 of the N-end rule
pathway.";
RL Hum. Mol. Genet. 13:2421-2430(2004).
RN [7]
RX VARIANTS ASP-267 AND GLN-1005.
RX PubMed=12601557; DOI=10.1007/s100380300016;
RA Roverati G., Beghini A., Zamboni G., Paradisi M., Larizza L.,
RT "Identification of two novel RECQL4 exonic SNPs and genomic
characterization of the IVS2 minisatellite.";
RL J. Hum. Genet. 48:107-109(2003).
RN [8]
RX VARIANTS GLN-355; SER-441 AND 857-CYS--THR-858 DEL.
RX PubMed=15221963; DOI=10.1002/jc.20269;
RA Nishijo K., Nakayama T., Aoyama T., Okamoto T., Ishibe T., Yasura K.,
RA Shima Y., Shibata K.R., Tsuboyama T., Nakamura T., Toguchida J.,
RT "Mutation analysis of the RECQL4 gene in sporadic osteosarcomas.";
RL Int. J. Cancer 111:367-372(2004).
RN [9]
RX FUNCTION: DNA-dependent ATPase. May modulate chromosome
segregation.
RN [10]
RX SUBUNIT: Interacts with UBR1 and UBR2.
RN [11]
RX TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
thymus and testis.
RN [12]
RX INDUCTION: Up-regulated in actively dividing cells.
RN [13]
RX DISBAS: Defects in RECQL4 are a cause of Rothmund-Thomson
syndrome (RTS) [MIM:268400]. A disease characterized by
dermatological features such as atrophy, pigmentation, and
telangiectasia and frequently accompanied by juvenile cataract,
saddle nose, congenital bone defects, disturbances of hair growth,
and hypogonadism.
RN [14]
RX DISBAS: Defects in RECQL4 are a cause of RAPADILINO syndrome
[MIM:266280]. A disease characterized by radial and patellar
aplasia or hypoplasia.
RN [15]
RX SIMILARITY: Belongs to the helicase family. RecQ subfamily.
RX DATAABAS: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
WWW="http://www.infobiochem.fr/services/chromancer/Genes/RECQL4ID285.html".
RN [16]
RX This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
RN [17]
RX EMBL, AB006532; BAA74453.1; -; mRNA.
RX EMBL, AB026546; BAA86899.1; -; Genomic DNA.
RX EMBL, BC011602; AAH11602.2; -; mRNA.
RX EMBL, BC013277; AAH13277.2; -; ALT_INIT; mRNA.
RX HSSP; P15043; 10Y.
RX Ensemble; ENSG00000160957; Homo sapiens.
RX HGNC; HGNC:9949; RECQL4.

DR MIM; 603780; -;
DR MIM; 268400; -;
DR MIM; 266280; -;
DR GO; GO:0003678; P:DNA helicase activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0006281; P:DNA repair; TAS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002464; DEAH box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR PANTHER; PTHR13710; RecQ; 1.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR TIGRPFAM; TIGR00614; recQ_fam; 1.
DR PROSITE; PS00690; DEAH_ATP_HELICASE; Nucleic acid binding; Polymorphism.
KW ATP-binding; Cataract; Helicase; Hydrolyase; Nuclear protein;
KW Nucleotide-binding; Polymorphism.
FT NE_BIND 502
FT MOTIF 605 608
FT VARIANT 267 267
FT VARIANT 355 355
FT VARIANT 441 441
FT VARIANT 857 858
FT VARIANT 1005 1005
SQ SEQUENCE 1208 AA; 133077 MW; CB809A7765AB48A1 CRC64;
Query Match 100.0%; Score 6424; DB 1; Length 1208;
Best Local Similarity 100.0%; Pred.No. 2.7e-307;
Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MERLRDVRERLQAMERAFRORGRPSODVEAPEETALYREYTLKRTTQAGGGR 60
DB 1 MERLRDVRERLQAMERAFRORGRPSODVEAPEETALYREYTLKRTTQAGGGR 60
QY 61 SSESIPAAAEAPRCPWPHNRAATKSPPTGRSROGSVDPYQRLKANIKGTLOAG 120
DB 61 SSESIPAAAEAPRCPWPHNRAATKSPPTGRSROGSVDPYQRLKANIKGTLOAG 120
QY 121 PALGRPPWPLGAASASTPKPGTVPVSPFAKVSDEPPQLEPPPPRGRLOHLQASIS 180
DB 121 PALGRPPWPLGAASASTPKPGTVPVSPFAKVSDEPPQLEPPPPRGRLOHLQASIS 180
QY 181 QRLGSIDPGLORCHSEVDFGAPACRPPDGSSESLIPGESAVLPGAGSGQPEAS 240
DB 181 QRLGSIDPGLORCHSEVDFGAPACRPPDGSSESLIPGESAVLPGAGSGQPEAS 240
QY 241 AFQEVSIKVSPPSSSGGKRRWNEEPWESPQVQOESSQAGPPSEGAGAVAEEDPPG 300
DB 241 AFQEVSIKVSPPSSSGGKRRWNEEPWESPQVQOESSQAGPPSEGAGAVAEEDPPG 300
QY 301 BEVQAPPPPCSSPNPRHGISPSQARAGAGETAPLHITPRLARHNRGNVRLNMQ 360
DB 301 BEVQAPPPPCSSPNPRHGISPSQARAGAGETAPLHITPRLARHNRGNVRLNMQ 360
QY 361 KHYVGRALRSRLRKQAKQKRGKGCFFGGGATVTTKSCFLENEOPDHMAAOCPRPA 420
DB 361 KHYVGRALRSRLRKQAKQKRGKGCFFGGGATVTTKSCFLENEOPDHMAAOCPRPA 420
QY 421 SEEDTDVAGPEPLVSPQVPEVPSLPTVLPYLSIGSGQLAETPAVFOALBQJGHOA 480
DB 421 SEEDTDVAGPEPLVSPQVPEVPSLPTVLPYLSIGSGQLAETPAVFOALBQJGHOA 480
QY 481 PRPGGRVAVMRLSGISTTLVPTGAGSICVQALALYSRSPCLTLYVSLTSLMDQ 540
DB 481 PRPGGRVAVMRLSGISTTLVPTGAGSICVQALALYSRSPCLTLYVSLTSLMDQ 540
QY 541 VSGLPCLKAACIHSGMTRKORBSVLOKIRAAQVHVMLTPEALVAGGLPPAAQLPVA 600

10

DB 541 VSGIPPLCKAKCHISGNTKQRESEVLQKIRAAQVHVALMTPEALVAGAGLPPAAGLPVA 600
|||
QY 601 PACIDEAHCHLSQMSHNFRPCYLRVCKVLRBMGVHCHCTGLTATVARTSTSDVAQHLAAVE 660
DB 601 PACIDEAHCHLSQMSHNFRPCYLRVCKVLRBMGVHCHCTGLTATVARTSTSDVAQHLAAVE 660
QY 661 EPDLHGPAVPTNTHLSVSDRDTQALLTLQGRFQMLDSIIYYCNREDTERIALLL 720
661 EPDLHGPAVPTNTHLSVSDRDTQALLTLQGRFQMLDSIIYYCNREDTERIALLL 720
QY 721 RTCLHAANVPSSGGAPKTTAAEYAHGMSRERRRVOARFMQGLRVYVATVAFGGLDR 780
721 RTCLHAANVPSSGGAPKTTAAEYAHGMSRERRRVOARFMQGLRVYVATVAFGGLDR 780
DB 721 RTCLHAANVPSSGGAPKTTAAEYAHGMSRERRRVOARFMQGLRVYVATVAFGGLDR 780
DB 781 PDVRAVLTGLTLPSPFESYVOAVAGAGDQPAHCHFLQPGQGLDRLRSHVHADSTDFL 840
|||
QY 841 AVKRLVQVRVPACTCTCTTPRPSBQEGAVGGRPVPKYPPQEAQQLSHQAAVGPBRVCMGH 900
841 AVKRLVQVRVPACTCTCTTPRPSBQEGAVGGRPVPKYPPQEAQQLSHQAAVGPBRVCMGH 900
DB 841 AVKRLVQVRVPACTCTCTTPRPSBQEGAVGGRPVPKYPPQEAQQLSHQAAVGPBRVCMGH 900
QY 901 ERALPLOTVOALDMPERALETLLCYLELHPHMLLELLATTYTCRLNCPGGPAQLQALA 960
|||
DB 901 ERALPLOTVOALDMPERALETLLCYLELHPHMLLELLATTYTCRLNCPGGPAQLQALA 960
QY 961 HRCPLAVCLAQQLPEDPGQSSSVFPMVNLVDSMGELASVRAQLCOLQMDHPRGV 1020
961 HRCPLAVCLAQQLPEDPGQSSSVFPMVNLVDSMGELASVRAQLCOLQMDHPRGV 1020
DB 961 HRCPLAVCLAQQLPEDPGQSSSVFPMVNLVDSMGELASVRAQLCOLQMDHPRGV 1020
QY 1021 RRGTVLVEFSELAFLHLSRQDLTLREKQDQCDPLVGRVQAREKQALRLRTFOAHSV 1080
1021 RRGTVLVEFSELAFLHLSRQDLTLREKQDQCDPLVGRVQAREKQALRLRTFOAHSV 1080
DB 1021 RRGTVLVEFSELAFLHLSRQDLTLREKQDQCDPLVGRVQAREKQALRLRTFOAHSV 1080
QY 1081 APFGSGPLCBODDERSTELKQILGRYPERESQBPGEKEDAQGPFGQALQMDMEDVR 1140
1081 APFGSGPLCBODDERSTELKQILGRYPERESQBPGEKEDAQGPFGQALQMDMEDVR 1140
DB 1081 APFGSGPLCBODDERSTELKQILGRYPERESQBPGEKEDAQGPFGQALQMDMEDVR 1140
QY 1141 CDIROPFLSRPEKTSRAVARIFHGIGSPCYVAQVYGGDRRFRWRKTLHLSPALVGLAT 1200
1141 CDIROPFLSRPEKTSRAVARIFHGIGSPCYVAQVYGGDRRFRWRKTLHLSPALVGLAT 1200
DB 1201 BELLOYAR 1208
1201 BELLOYAR 1208
|||
DB 1201 BELLOYAR 1208

RESULT 2
RECO4 MOUSE STANDARD; PRT: 1216 AA.
ID RECO4 MOUSE
AC 075NR7; Q76MT1; Q99PV9;
DT 13-SRP-2005 (Rel. 48, Created)
DT 13-SRP-2005 (Rel. 48, Last sequence update)
DT 13-SRP-2005 (Rel. 48, Last annotation update)
DB ATP-dependent DNA helicase Q4 (BC 3.6.1.-) (RecQ protein-like 4).
DN Name=Recq4; Synonym=Recq4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN [1]
RN NUCLEOTIDE SEQUENCE (GENOMIC DNA / MRNA) (ISOFORM 1) AND NUCLEOTIDE
RP SEQUENCE (MRNA) OF 16-1115 (ISOFORM 2).
RX PubMed=1167012;
RA Ohnata T., Araki R., Fukumura R., Kuroiwa A., Matsuda Y., Tatsuami K.,
RA Abe M.;
RT "Cloning, genomic structure and chromosomal localization of the gene
RL encoding mouse DNA helicase Recq protein-like 4."
RN Gene 261:251-258(2000).
RP [2]
FUNCTION.

| | |
|----|--|
| XX | PubMed=12915449; DOI=10.1093/hmg/ddg254; |
| RA | Hoki Y., Arai R., Fujimori A., Ohnata T., Koseki H., Fukumura R., |
| RA | Nakamura M., Takahashi H., Noda Y., Kito S., Abe M.; |
| RA | "Growth retardation and skin abnormalities of the Recq14-deficient |
| RT | mouse."; |
| RT | Hum. Mol. Genet. 12:2293-2299(2003). |
| RN | [3] |
| RX | DEVELOPMENTAL STAGE. |
| RP | PubMed=12952869; DOI=10.1093/hmg/ddg306; |
| RA | Siltonen H.A., Kopra O., Kaeerlahti J., Haravuori H., Winter R.M., |
| RA | Saarenen A.-M., Peltonen L., Kastl M.; |
| RA | "Molecular defect of RAPADILINO syndrome expands the phenotype |
| RT | spectrum of RCOL diseases."; |
| RL | Hum. Mol. Genet. 12:2837-2844(2003). |
| RN | [4] |
| RP | FUNCTION. |
| RA | PubMed=15703196; DOI=10.1093/hmg/ddi075; |
| RX | Mann M.B., Hodges C.A., Barnes B., Vogel H., Haasold T.J., Luo G.; |
| RT | "Defective sister-chromatid cohesion, aneuploidy and cancer |
| RT | predisposition in a mouse model of type II Rothmund-Thomson |
| RT | syndrome."; |
| RL | Hum. Mol. Genet. 14:813-825(2005). |
| CC | -1- FUNCTION: DNA-dependent ATPase (By similarity). May play a role in |
| CC | development of the palate and the limbs. May modulate chromosome |
| CC | segregation. |
| CC | -1- SUBUNIT: Interacts with UBR1 and UBR2 (By similarity). |
| CC | -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity). |
| CC | -1- ALTERNATIVE PRODUCTS: |
| CC | Bvent-Alternative splicing; Named isoforms=2; |
| CC | Name=1; |
| CC | IsoId=Q75NR7-1; Sequence=Displayed; |
| CC | Name=2; |
| CC | IsoId=Q75NR7-2; Sequence=VSP_015177; |
| CC | -1- DEVELOPMENTAL STAGE: Not expressed at E12.5. Expressed at E15.5- |
| CC | E18.5, with highest levels in chondrocytes of developing bone and |
| CC | cartilage and immature proliferating enterocytes of intestine. |
| CC | -1- MISCELLANEOUS: Absence of the protein causes early embryonic |
| CC | lethality. Transgenic mice with exon 13-deleted RCOL4 are severely |
| CC | growth-retarded and show high (98%) perinatal lethality. They |
| CC | exhibit various skin, bone, intestine, tooth and thymus |
| CC | abnormalities and premature aging features, but have normal |
| CC | sensitivity to IR and UV irradiation. In contrast, transgenic mice |
| CC | expressing a truncated form of RCOL4 exhibit mild perinatal |
| CC | lethality, no growth defect, but show defects of the skin and |
| CC | skeleton, aneuploidy and increased cancer susceptibility. |
| CC | -1- SIMILARITY: Belongs to the helicase family. Recq subfamily. |
| CC | -1- SIMILARITY: Contains 1 CCHC-type zinc finger. |
| CC | ----- |
| CC | This Swiss-prot entry is copyright. It is produced through a collaboration |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - |
| CC | the European Bioinformatics Institute. There are no restrictions on its |
| CC | use as long as its content is in no way modified and this statement is not |
| CC | removed. |
| CC | ----- |
| DR | EMBL; AB039882; BAD1131.1; -; mRNA. |
| DR | EMBL; AB175741; BAD14289.1; -; mRNA. |
| DR | EMBL; AB042529; BAB32696.1; -; genomic_DNA. |
| DR | HSSP; P15043; IOTY. |
| DR | Ensembl; ENSMUSG0000033762; Mus musculus. |
| DR | MGI; MGI:1931028; Recq14. |
| DR | PROSITE; PS00690; DEAH ATP HELICASE, FALSE_NEG. |
| DR | PROSITE; PS01581; ZF_CCHC; 1. |
| KW | Alternative splicing; ATP-binding; Helicase; Hydrolase; Metal-binding; |
| KW | Nuclear protein; Nucleotide-binding; Zinc; Zinc-finger. |
| KW | CCHC-type. |
| KW | ATP (potential). |
| KW | DEAH box. |
| FT | NP BIND 393 410 |
| FT | ZN BIND 519 526 |
| FT | MOTIF 637 630 |
| FT | VARSPLIC 843 843 |
| FT | CONFLICT 732 739 |
| FT | CONFLICT 878 878 |
| FT | CONFIDENCE 878 878 |
| FT | S -> G (in Ref. 1; BAD11311). |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:54:12 ; Search time 12 Seconds
(without alignments)
718.036 Million cell updates/sec

Title: US-09-889-325-4
Perfect score: 6424
Sequence: 1 MEKLRDVRERLQWERAFFR.....HLSFHALVGLATELLOVAR 1208

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 54001 seqs, 7132810 residues

Total number of hits satisfying chosen parameters: 54001

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB pep.*
5: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB pep.*
7: /cgn2_6/ptodata/2/pubaa/US11_NEW_PUB pep.*
8: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 481 | 7.5 | 767 | 6 | US-10-467-657-2430 |
| 2 | 481 | 7.5 | 767 | 6 | US-10-467-657-2430 |
| 3 | 476 | 7.4 | 1405 | 6 | US-10-995-561-529 |
| 4 | 476 | 7.4 | 1406 | 6 | US-10-995-561-530 |
| 5 | 476 | 7.4 | 1436 | 6 | US-10-995-561-531 |
| 6 | 232.5 | 3.6 | 1823 | 6 | US-10-995-561-988 |
| 7 | 232.5 | 3.6 | 2102 | 6 | US-10-995-561-990 |
| 8 | 232.5 | 3.6 | 2157 | 6 | US-10-995-561-991 |
| 9 | 229 | 3.6 | 2108 | 6 | US-10-995-561-989 |
| 10 | 225.5 | 3.5 | 135 | 6 | US-10-793-626-3134 |
| 11 | 224 | 3.5 | 411 | 6 | US-10-793-626-3156 |
| 12 | 207.5 | 3.2 | 614 | 6 | US-10-519-447-2 |
| 13 | 207.5 | 3.2 | 614 | 6 | US-10-878-556A-75 |
| 14 | 199 | 3.1 | 882 | 7 | US-11-012-762-34 |
| 15 | 194.5 | 3.0 | 650 | 6 | US-10-878-556A-110 |
| 16 | 191.5 | 3.0 | 457 | 6 | US-10-467-657-2172 |
| 17 | 170.5 | 2.7 | 874 | 7 | US-11-012-762-8 |
| 18 | 169.5 | 2.6 | 411 | 6 | US-10-821-234-1386 |
| 19 | 169.5 | 2.6 | 411 | 6 | US-10-878-556A-140 |
| 20 | 166 | 2.6 | 1560 | 7 | US-11-059-982-1 |
| 21 | 165 | 2.6 | 406 | 6 | US-10-878-556A-41 |
| 22 | 163 | 2.5 | 915 | 6 | US-10-995-561-1003 |
| 23 | 163 | 2.5 | 940 | 6 | US-10-995-561-1004 |
| 24 | 163 | 2.5 | 969 | 6 | US-10-995-561-1001 |
| 25 | 163 | 2.5 | 994 | 6 | US-10-995-561-997 |

| | | | | | |
|----|-------|-----|------|---|--------------------|
| 26 | 158.5 | 2.5 | 917 | 6 | US-10-995-561-1000 |
| 27 | 158.5 | 2.5 | 971 | 6 | US-10-995-561-998 |
| 28 | 157.5 | 2.5 | 680 | 6 | US-10-467-657-2008 |
| 29 | 157 | 2.4 | 628 | 6 | US-10-995-561-1002 |
| 30 | 157 | 2.4 | 740 | 6 | US-10-821-234-1464 |
| 31 | 153 | 2.4 | 915 | 6 | US-10-821-234-1514 |
| 32 | 150.5 | 2.3 | 674 | 6 | US-10-821-234-965 |
| 33 | 149.5 | 2.3 | 428 | 7 | US-11-000-463-448 |
| 34 | 149 | 2.3 | 462 | 6 | US-10-467-657-4348 |
| 35 | 148.5 | 2.3 | 1133 | 6 | US-10-821-234-1219 |
| 36 | 148.5 | 2.3 | 2630 | 7 | US-11-186-731-2 |
| 37 | 148.5 | 2.3 | 7968 | 7 | US-11-186-731-5 |
| 38 | 147.5 | 2.3 | 1377 | 6 | US-10-821-234-1070 |
| 39 | 146.5 | 2.3 | 660 | 6 | US-10-878-556A-102 |
| 40 | 145.5 | 2.3 | 495 | 7 | US-11-074-176-266 |
| 41 | 144.5 | 2.2 | 615 | 6 | US-10-982-545-14 |
| 42 | 142 | 2.2 | 3073 | 7 | US-11-143-980-50 |
| 43 | 139.5 | 2.2 | 447 | 6 | US-10-995-561-930 |
| 44 | 138.5 | 2.2 | 980 | 7 | US-11-064-246-10 |
| 45 | 138 | 2.1 | 1618 | 6 | US-10-984-645-2 |

ALIGNMENTS

RESULT 1
US-10-467-657-2430
; Sequence 2430, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2430
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2430

| | | | | |
|-----------------------|------------------|--|------------|-------------|
| Query Match | 7.5% | Score 481; | DB 6; | Length 767; |
| Best Local Similarity | 33.1% | Pred. No. 7.1e-24; | | |
| Matches 139; | Conservative 64; | Mismatches 169; | Indels 48; | Gaps 13; |
| QY | 474 | EQLGHQAFRCQERAVMRILSGISTLLVLP TGACKSLCYQLPALLYSRSPCLTLWVSP 533 | | |
| Db | 13 | EVGYPERGRQEDVINTLAGGSLTVLMPFGGKSLCYQIPALMREG-----VAVVWSPL 68 | | |
| QY | 534 | LSLMDQVSGLPPOC-LKAACIHSQWTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPP 592 | | |
| Db | 69 | IALMNDQVASLHVAGIEAAVNSGTSADAEARIDAKLQGRLLKLYVAPERLVTDRFLRF 128 | | |
| QY | 593 | AAQLPPVAFACIDRAHCLSQSHNFRPCYLKVKVLRRMGVHFLGLTATATRTASDV 652 | | |
| Db | 129 | LDQQTVSLIFA-IDEAHCVSRWGHDFRPEYQQL-GMLAERYENIPRIALTATADAATRA 186 | | |
| QY | 653 | AQLHVAEEPLDHCAPVPTNLHLSVSMRDRTDQALLTLQGRFQNLDSIIIVCNRED 712 | | |
| Db | 187 | KHYLHLDQASEFVSFD-RTNIYYQVTEKNNGKQLDIFI--RKEMTQSGIVYCLSRK 243 | | |
| QY | 713 | TERIAALLRTCLHAAVPGSGGRAPKTTAAEYHAGMCSRERRRRVQRAFMOGLRVVAVT 772 | | |
| Db | 244 | VEDAAQFLR-----ENG-----LNAIPYHAGLSMDVRENGQRRFTTHEDNIIVAVT 289 | | |

[illegible]

```

RESULT 5
US-10-995-561-531
; Sequence 531, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 531
; LENGTH: 1436
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-531

```

```
Qy 1050 QICDPLYGR-VOAREROA 1066
Db 1157 QI--VLXGKLYEARQKHA 1172

RESULT 6
US-10-995-561-988
; Sequence 988, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 988
; LENGTH: 1823
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-988

Query Match 3.6%; Score 232.5; DB 6; Length 1823;
Best Local Similarity 22.6%; Pred. No. 1.7e-07;
Matches 222; Conservative 72; Mismatches 325; Indels 363; Gaps 47;

Qy 15 ERAPRRQRGRPPSODDVEAAPEETRALYREYRTLKRTTGQAGGLRSESLSLPA---EE 71
Db 904 ERPRRRHGRQAQOD-----KPPFRRLKQE-----RENAARGSEKPSLTLPASAPGPEE 954
Qy 72 -----APEPRCWGPHLNRAATKSPQ-----
Db 955 ALTTVTVAAPR-----RAAAKSPDLNSQNSQDANEWEETASESSDFTSERRGDKPEAP 1007
Qy 92 -----PTGRSRQGSVP-----DYQRLKANLKGTLQA-CPAL---GRPPWP 129
Db 1008 PPVLLTPKAVGTPGGGGGAVPGISAMSRGDLSQLAKDLKRSFSSQRPQGMQRNRRPGP 1067
Qy 130 LGRASSKASTPKPGTGP-----VPSFAEKVS-----DEPPOLPEQP----- 167
Db 1068 GKGAGSSGSSGGGGGPGGRTGCGDKRSWPKNRSRPPKRSRPPKRSRPPKRSRPPKRSRPP 1127
Qy 168 -----RPGRLQHLQASLSQRLGSLD-PGWLQR-----
Db 1128 VFRLDQVIHNSNPAGIQALASLSRQGSVTPAGGHPKRPKPPQAPQGPSPRPTRVEPQ 1187
Qy 194 -----CHSEVPDFL-----GAPKAC-----RDLGSEESQLLIPG 223
Db 1188 RVNSGLSSDPHFEPGPMVRGVGTGTPRDSAGVSPFPKRRERPRKPELLOEES-LPPPH 1246
Qy 224 ESAVLGAGAGQGEAFQVSVTRVQSP-----QPSGSGEYRWNNEPWESE----- 272
Db 1247 SSGFLGSKPEGPGQABES-RDTGTEALTPHWNRLHTATSKSRYPSSMBPMEPLSPFFE 1305
Qy 273 -----AQV-----QOESS-----QAGPPSSEG-----AVAVE 295
Db 1306 DVAGTEMSQSDSGVDLSGDSQVSSGPCSQRSPPDGLKGAABGPKRPKPGSSPLNAPVCE 1365
Qy 296 EDPGGEVQAPQPCSSPSN--PRYHGLSPS--SQARAGKAE-GTAPLHIFPRLARHDR 350
Db 1366 GPPGSEPPRPAPPHGDKELREQLPPLPGPIGTERSQTDRTGTFGPIRP---SHRP 1422
Qy 351 GNYVRLNMKQ-----HYVRGALRSRLRLKQAWKQKWKKBCEGCGGATVTKESCFLN 406
Db 1423 GPPYQFGTSKDSLDRLVWGDLSK-----AEKE-----LTASVTEAIPVS 1462
Qy 407 EQDF--HWAQCPAPSE--DTDVAGPEPLVPSQPVPVEPSLDPTVLPLY-SLGPSSQL 462
Db 1463 RDWELLSAASAPQSKNLDSCHCVEPSSSGQRLYPEV-----PYGSAGPSSQ 1513
```

```
Qy 463 AETPAEVFQALEQLGHQAFRPGQ-----ERAVNRILSGI-----S 497
Db 1514 ISGAMDSQLHPNSG--GFRFGTSLHPYRSQPLYLPPGPAPPALLSGVALKGQFLDFS 1571
Qy 498 TLIV-----LPTGAGKSLCYQLPALLYSR--SP-----CLTLVVSPLLSLMDQVSGLP 545
Db 1572 TMOATELGKLPAG--GVLYPPSPFLYSPAFCSPLPDTSLQVRODLSPSDFYSTPLQ 1628
Qy 546 PCLKAAACIHSGMTRKQRESVLQIRAAQVHVLMLTPEA-----LVGAGGLPPAAQLPVA 600
Db 1629 PGSGSGFLPSGAPAQ-----MLLPWDSQLPVNFGSLPPAPPPPPP 1672
Qy 601 FACIDEAHCLSQSHNFRPCYLKVLRRMGVHCFGLGTATATRTTASDVAQHLAAVE 660
Db 1673 LSLLPVGPALQPPSLAVRP-----PPAPATRVLPSPARPPFASLG 1712
Qy 661 EPDLHGPAVPTNHLVSMDRDTQALLTLLOKRRFQNLDSII-----IYCNRE 711
Db 1713 RAEHLPVVELKPFQDYKLLSNLGGPGSSRTPPTGRSFGSLNSRLKATFTSYGVFTQRV 1772
Qy 712 DTERIAA---LLRTCLHAAMP 730
Db 1773 DLYQASPPDALR-----WIP 1788

RESULT 7
US-10-995-561-990
; Sequence 990, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 990
; LENGTH: 2102
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-990
```

```
Query Match 3.6%; Score 232.5; DB 6; Length 2102;
Best Local Similarity 22.6%; Pred. No. 2e-07;
Matches 222; Conservative 72; Mismatches 325; Indels 363; Gaps 47;

Qy 15 ERAPRRQRGRPPSODDVEAAPEETRALYREYRTLKRTTGQAGGLRSESLSLPA---EE 71
Db 1183 ERPRRRHGRQAQOD-----KPPFRRLKQE-----RENAARGSEKPSLTLPASAPGPEE 1233
Qy 72 -----APEPRCWGPHLNRAATKSPQ-----
Db 1234 ALTTVTVAAPR-----RAAAKSPDLNSQNSQDANEWEETASESSDFTSERRGDKPEAP 1286
Qy 92 -----PTGRSRQGSVP-----DYQRLKANLKGTLQA-CPAL---GRPPWP 129
Db 1287 PPVLLTPKAVGTPGGGGGAVPGISAMSRGDLSQLAKDLKRSFSSQRPQGMQRNRRPGP 1346
Qy 130 LGRASSKASTPKPGTGP-----VPSFAEKVS-----DEPPOLPEQP----- 167
Db 1347 GKGAGSSGSSGGGGGPGGRTGCGDKRSWPKNRSRPPKRSRPPKRSRPPKRSRPPKRSRPP 1406
Qy 168 -----RPGRLQHLQASLSQRLGSLD-PGWLQR-----
Db 1407 VFRLDQVIHNSNPAGIQALASLSRQGSVTPAGGHPKRPKPPQAPQGPSPRPTRVEPQ 1466
Qy 194 -----CHSEVPDFL-----GAPKAC-----RDLGSEESQLLIPG 223
Db 1467 RVNSGLSSDPHFEPGPMVRGVGTGTPRDSAGVSPFPKRRERPRKPELLOEES-LPPPH 1525
```

```

Qy 224 ESAVLGAGAGSOGPEASAFQEVSVIRVSGP-----QPSSSGGKKRWNEEPWESP----- 272
Db 1526 SSGFLGSKPEGPGQAES-RDTGTEALTTHIWNRLHTATSRKSYRPSMEPMEPLSPFE 1584
Qy 273 -----AQV-----QOESS-----QAGPPSEAG-----AVAVE 295
Db 1585 DVAGTEMSQSDSGVDLSGDSQVSSGPCSQRSPDGLKGAAEGPKRPGSGSPLNAVPC 1644
Qy 296 EDPGEPVQAPPOPCSSPSN--PRYHGLSPS--SQARAGKAE--GTAPLHIFPLARHDR 350
Db 1645 GPPGSEPPRRPPAPHDGRKELPREQPLPPGPIGTERTSORDTGTGTPGPIRP---SHRP 1701
Qy 351 GNYVRLNKKQK----HYVRGALRSRLLRKQAWKQKWKKECGFCGGGATVTTKESCFLN 406
Db 1702 GPPVQFGTSDKSDRLRLVVGDSLK-----AEKE-----LTASVTEAIPVS 1741
Qy 407 EQFD--HWAACQPRPASEE-DTDAVGPEPLVPSQPVPVEPSLDPTVLPLY-SLQPSGQL 462
Db 1742 RDWELLPSAASAEFQSKNLDGCHVPEPSSGGQRLYPEV-----FYGSAGPSSQ 1792
Qy 463 AETPAEVQFQALQGHQAFRPGQ-----QOESS-----ERAVMRILSGI-----S 497
Db 1793 ISGAMDSQLHPNSG--GFRGTPSLHPYRSQPLYLPPGPAPPALLSGVALKGQFLDFS 1850
Qy 498 TLLV-----LPTGAGKSLCYQLPALLYRR--SP-----CLTLVVSPLLSLMDDOVSLP 545
Db 1851 TMQATELGKLPAG---GVLYPPPSFLYSPAFPCSPPLDTSLLQVRQDLPSDFYSTPLQ 1907
Qy 546 PCLKAACTHSGWTRKQRESVLQKIRAAQVHVLMLTPEA-----LVGAGGLPAAQLPPVA 600
Db 1908 PGGSGGLFSPGAPAQO-----MLLPWVDSQLPVNFSGSLPPAPPAPP 1951
Qy 601 PACIDEAHLCSQWSHNFRPCYLVRCKVLRERMGVHCFGLGTATATRTTASDVAAHLVAE 660
Db 1952 LSLPLVPGALQPPSLAVRP-----PPAPATRVLPSPARPPASLG 1991
Qy 661 EPDLHGAPVPTNLHLSVMDRDTQALLTLQKRFQNLDSII-----IYCNRE 711
Db 1992 RAEHLFVELKPFQDYQKLSNLGGPSSRTPTGSRFSGLSRLKATPTSTYSGVFTQRV 2051
Qy 712 DTERIAA---LLRTCLHAAPV 730
Db 2052 DLYQOASPPDALR-----WIP 2067

RESULT 8
US-10-995-561-991
; Sequence 991, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 991
; LENGTH: 2157
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-991

```

```

Query Match 3.6%; Score 232.5; DB 6; Length 2157;
Best Local Similarity 22.6%; Pred. No. 2.1e-07;
Matches 222; Conservative 72; Mismatches 325; Indels 363; Gaps 47;

Qy 15 ERAFRQRGRSODDVEAAPEETRALYREYRTLKTTGQAGGLRSSLSLPA---EE 71
Db 1238 ERPRRRHGAQQD-----KPRFRRLKQF-----RENARSGKPSLTLPASAPGEE 1288

```

```

Qy 72 -----APEPRCWMGPHLNRATKSPQ----- 91
Db 1289 ALTTVTVPAPR-----RAAKSPDLNSQNSQANEWEETASSSDFTSERRGKEAP 1341
Qy 92 -----PTGSRROGSVP-----DYQORLKANLKGTLQA--GPAL---GRPPWP 129
Db 1342 PPVLLTPKAVGTTPGGGGGAVPGISAMSRGDLGSRANDLSKRFSQRPQMERQNRPPG 1401
Qy 130 LGRASSKASTPKPGTGP---VPSFAEKYS-----DEPQLPEPQ--- 167
Db 1402 GKGAGSGSGSGGGGPGGRTGFGCDKRSWSPKMRSPRPPPEPGLPLPPPPSSA 1461
Qy 168 -----RPGRLQHLQASLSORLSLD--PGWLQR----- 193
Db 1462 VFRLDQVIHNAPAGIQALQALSRSQSVTAPGCHPHKPGPQAPQPGSPRPTRYEPQ 1521
Qy 194 -----CHSEVPDFL-----GAPKAC-----RPDLSEESQLLIPG 223
Db 1522 RVNSGLSSDPHFEEPGVMVRGVGTPRDSAGVSPFPKRRRPPRKPPELLQEE--LPPPH 1580
Qy 224 ESAVLGAGAGSOGPEASAFQEVSVIRVSGP-----QPSSSGGKKRWNEEPWESP----- 272
Db 1581 SSGFLGSKPEGPGQAES-RDTGTEALTTHIWNRLHTATSRKSYRPSMEPMEPLSPFE 1639
Qy 273 -----AQV-----QOESS-----QAGPPSEAG-----AVAVE 295
Db 1640 DVAGTEMSQSDSGVDLSGDSQVSSGPCSQRSPDGLKGAAEGPKRPGSGSPLNAVPC 1699
Qy 296 EDPGEPVQAPPOPCSSPSN--PRYHGLSPS--SQARAGKAE--GTAPLHIFPLARHDR 350
Db 1700 GPPGSEPPRRPPAPHDGRKELPREQPLPPGPIGTERTSORDTGTGTPGPIRP---SHRP 1756
Qy 351 GNYVRLNKKQK----HYVRGALRSRLLRKQAWKQKWKKECGFCGGGATVTTKESCFLN 406
Db 1757 GPPVQFGTSDKSDRLRLVVGDSLK-----AEKE-----LTASVTEAIPVS 1796
Qy 407 EQFD--HWAACQPRPASEE-DTDAVGPEPLVPSQPVPVEPSLDPTVLPLY-SLQPSGQL 462
Db 1797 RDWELLPSAASAEFQSKNLDGCHVPEPSSGGQRLYPEV-----FYGSAGPSSQ 1847
Qy 463 AETPAEVQFQALQGHQAFRPGQ-----QOESS-----ERAVMRILSGI-----S 497
Db 1848 ISGAMDSQLHPNSG--GFRGTPSLHPYRSQPLYLPPGPAPPALLSGVALKGQFLDFS 1905
Qy 498 TLLV-----LPTGAGKSLCYQLPALLYRR--SP-----CLTLVVSPLLSLMDDOVSLP 545
Db 1906 TMQATELGKLPAG---GVLYPPPSFLYSPAFPCSPPLDTSLLQVRQDLPSDFYSTPLQ 1962
Qy 546 PCLKAACTHSGWTRKQRESVLQKIRAAQVHVLMLTPEA-----LVGAGGLPAAQLPPVA 600
Db 1963 PGGSGGLFSPGAPAQO-----MLLPWVDSQLPVNFSGSLPPAPPAPP 2006
Qy 601 PACIDEAHLCSQWSHNFRPCYLVRCKVLRERMGVHCFGLGTATATRTTASDVAAHLVAE 660
Db 2007 LSLPLVPGALQPPSLAVRP-----PPAPATRVLPSPARPPASLG 2046
Qy 661 EPDLHGAPVPTNLHLSVMDRDTQALLTLQKRFQNLDSII-----IYCNRE 711
Db 2047 RAEHLFVELKPFQDYQKLSNLGGPSSRTPTGSRFSGLSRLKATPTSTYSGVFTQRV 2106
Qy 712 DTERIAA---LLRTCLHAAPV 730
Db 2107 DLYQOASPPDALR-----WIP 2122

```

```

RESULT 9
US-10-995-561-989
; Sequence 989, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

```

```
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 989
; LENGTH: 2108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-989

Query Match
Best Local Similarity 3.6%; Score 229; DB 6; Length 2108;
Matches 198; Conservative 56; Mismatches 258; Indels 328; Gaps 44;

QY 15 ERAPRRGRGRPSDDVEAAPEETRALVREYRTLKRTTGQAGGCLRSESELPAAA---EE 71
Db 1183 ERPRRRHGRAQQD-----KPRFRRLKQE-----RENAAGSGKPSLTPASAPGPEE 1233
QY 72 -----APRRCWGPHLNRAATKSPQ-----
Db 1234 ALTTVTVPAPR-----RAAAKSPDLNQNQSDQANEENETASESSDFTSERRGDKAP 1286
QY 92 -----PTPGRSROGVP-----DYQRLKANLKGTLQA-GPAL---GRRPWP 129
Db 1287 PVLLTPKAVGTGPGGGGAVGFGISAMRSGDLSQAKDLKSRFSQRPQMERQNRFRGP 1346
QY 130 LGRASSKASTPKPGTGP---VPSFAEKVS-----DEPQLPEPQP----- 167
Db 1347 GKGAGSGSSGGGGGPGGTGPRGDKRWSPPKRSRPPERRPGLPLPPPPSSSA 1406
QY 168 -----RPGRLQHLQASLSQRLSLD-PGWLQR-----
Db 1407 VFRLDQVTHSNFAGIQALASLRSQSVTAAGPHRKPGPPQAPQPGSPRPTRYEPQ 1466
QY 194 -----CHSEVPDFL-----CAPKAC-----RDLGSEESQLLIPG 223
Db 1467 RVNSGLSDPHFEPPGPMVRGVGTGTRDSAGVSPFPKRRRRPRPKPELLQEEES-LPPPH 1525
QY 224 ESAVLPGAGSQGPEASAFQSVSRVSGP-----QPSSSGGEKKRWNERPWPESP----- 272
Db 1526 SSGFLGSKPEPGHQAES-RDTGTEALTPIHWNRLHTATSKSVTRPSMFMWEPMLSPFE 1584
QY 273 -----AQV-----QSESS-----QAGPPSEAG-----AVAVE 295
Db 1585 DVAGTEMSSQSDGVDLSDGSQVSSGPCSQRSPPDGLKGAAGPPKPPGSSPLNAVPC 1644
QY 296 EDPGEPVQAQPPQPCSSPSN--PRYHGLSPS--SQARAGKAE-GTAPLHIFPRLARHDR 350
Db 1645 GPPGSEPPRRPPAPHDGDRKELREQPLPFGPIGTERTSQRDTGRTEPFGPIRP---SHRP 1701
QY 351 GNYVRLNMKOK---HYVRGRALSRLLRKAQWKQKWKKEGCGGAGTATTKESCFLN 406
Db 1702 GPPVQFGTSKDSLDRLVVGDSLK-----AEKE-----LTASVTTEALPVS 1741
QY 407 EQFD--HMAAQCPRASEE-DTDAVGPEPLVPSPQVPVEPSLDPTVLPY-SLGPSSGQL 462
Db 1742 RDWELLPSAAASAEFPQSKNLDGHCVEPSSSGQRLYPEV-----FVSGAGPSSQ 1792
QY 463 AETPAEVFOALEQLGHQAFRPQ-----
Db 1793 ISGGMDSQLHPNSG--GFRPGTSLPHYRSQPLYPGPPAPPSPALLSGVALKGQFLDFS 1850
QY 498 TLLV-----LPTGAGKSLCYOLPALLYRR--SP-----CLTLVVSPLLSLMDQVSLP 545
Db 1851 TMOATELGKLPAG--GVLTPPPSFLYSAPFCPSPLPDTSLQVRQDLPSFSDYSPYTPQ 1907
QY 546 PCLKAACHSGMTKQRESVLQKIRAAQVHVLMTPEA-----LVGAGGLPFPAAQLPPVA 600
Db 1908 FGGQSGFLPSGAPAAQ-----MLLPWVDSOLPVPVNFSGSLPPA---PPPA 1948
```

```
RESULT 10
US-10-793-626-3334
; Sequence 3334, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3334
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3334

Query Match
Best Local Similarity 3.5%; Score 225.5; DB 6; Length 135;
Matches 53; Conservative 30; Mismatches 47; Indels 13; Gaps 4;

QY 462 LASTPAEVFOALQQLGHQAERPGQERAVMRILSGISTLLVLTGAGKSLCYOLPALLYSR 521
Db 1 MQETLSHYF-----GYKSPRQGEIITKILNHOHTLGVLTGGGKSICYVPELMOGG 54
QY 522 RSPCLTLVAVSPLSLMDMDQVSGLP--LKAACHSGMTKQRESVLQKIRAAQVHVLMLT 580
Db 55 ----TTIVISPLISLKMKDQVDQLQAMGIQAAYLNSSLTHKQKEIEBQIKRGALQFLVA 110
QY 581 PEALVGAG--GLPPAAQLPPVAF 601
Db 111 PERFENTFFLNLLRKIEIPLAF 133

RESULT 11
US-10-793-626-3156
; Sequence 3156, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3156
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3156

Query Match
Best Local Similarity 3.5%; Score 224; DB 6; Length 411;
Matches 48; Conservative 19; Mismatches 43; Indels 14; Gaps 2;

QY 704 IYCNREDTERIAALRTCLHAAWPGSGGRAPKTTAEAYHAGMCSRRRRRVQRAFMQG 763
Db 48 IYCVSTRKQVEE-----LHEAL-----NSEKIKSTIYHAGLTNKRIEAQNDFLYD 93
QY 764 QLAVVATVAFGMLRDPDVRVHLGLPSPFSYVQAVCRAGDQGPACHLFLQPGQE 823
```

Db 94 RVEVVIATNAPGMDKSNRVYHYNMPGDLSEYYQAGRAGRDGLKSECLILFSERDK 153
 Qy 824 DLRE 827
 Db 154 GLHE 157

RESULT 12
 US-10-519-447-2
 ; Sequence 2, Application US/10519447
 ; Publication No. US20050244829A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
 ; APPLICANT: Makoto OGINO
 ; APPLICANT: HIDEKI ENDOH
 ; TITLE OF INVENTION: METHOD FOR SCREENING AN AGENT FOR IMPROVING INSULIN RESISTANCE
 ; FILE REFERENCE: Q85576
 ; CURRENT APPLICATION NUMBER: US/10/519,447
 ; CURRENT FILING DATE: 2004-12-30
 ; PRIOR APPLICATION NUMBER: PCT/JP03/08367
 ; PRIOR FILING DATE: 2003-07-01
 ; PRIOR APPLICATION NUMBER: JP 2002-193814
 ; PRIOR FILING DATE: 2002-07-02
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 614
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-519-447-2

Query Match 3.2%; Score 207.5; DB 6; Length 614;
 Best Local Similarity 22.6%; Pred. No. 1.9e-06;
 Matches 110; Conservative 63; Mismatches 222; Indels 91; Gaps 19;
 Qy 387 GECFGGGATVTK-----ESCFLNQFDHAAQCPRASEEDTAVGPEPLVPS 436
 Db 31 GKKEGNGEKLKKNLDELKPEKNFYQHPD-----LARTAQEVETYSRSEKITVRG 86
 Qy 437 ---POPVEPVSLEDTVLYSLGSGQLAETPAEVFOALBQLGHQAFRPGQERAVMRIL 493
 Db 87 HNCPE-----VLFNFE-----ANFPANVMDVIARQNFTPTAIQAQGPVAL 129
 Qy 494 SGISTLLVPTGAGKSLCYQLPALLYRRSP-----CLTLVVSPLSLMDQVSG- 543
 Db 130 SGLDMVGVAQTGSGKTLSPALLPAIHHNQPFLEGRDGPICLVLAFTRELAQQVQVAAE 189
 Qy 544 -LPPC-LKAACTHSGMTRKQRESVLQIRAAQVHVLMTPEALVG--AGGLPPAAQLPP 598
 Db 190 YCRACRLKSTCIYGAPKGPQIRDLER---GVEICATPGRLIDFLECG---KTNLRR 241
 Qy 599 VAFACIDEAHCLSQSHNFRPCYLKRVCKVLRMGVHCFGLGTATATRTASDVAQ---H 655
 Db 242 TTYLVLDADRMLDM--GFEPQIRKIVDQIRPDRQTLKMSATWPKVEVQLAEDFLKDYIH 299
 Qy 656 LAVAEEDPLHGPAPVPTNLHL-----SVSMRDRTDQALLTLQKRFQNLDSIIICNRR 711
 Db 300 INI-----GALELSANHNLIQIVDVCHDVEKDEKLRLMEEIMSEKENKTIVEFV--E 349
 Qy 712 DTERIAALLRTCLHAAPVPGSGRAPKTTAEAYHAGMCSRERRRVRQAFMQGLRVVAT 771
 Db 350 TKRRCDLTRKWRDGM--PAMG-----IHGDKSQQERDQVNLNFKHGKAPILIAT 398
 Qy 772 VAFGMGLDRDPVRAVLHGLPPSFESYVQAVGRAGRDGQPAHCHLFLOP-----QGEDLRE 827
 Db 399 DVASGLDVEDVKFVINYDYPNSSEDIYHRIQRTARSTKTGTAYTFTFPNNIKQVSDLS 458
 Qy 828 LRRHVH 833
 Db 459 VLREAN 464
 RESULT 13
 US-11-012-762-34
 ; Sequence 34, Application US/11012762
 ; Publication No. US20050244815A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Georgia State University Research Foundation, Inc.
 ; TITLE OF INVENTION: Compositions and Methods for Viral Resistance Genes
 ; FILE REFERENCE: GSU1.PCT
 ; CURRENT APPLICATION NUMBER: US/11/012,762
 ; CURRENT FILING DATE: 2004-12-15
 ; PRIOR APPLICATION NUMBER: PCT/US03/19300
 ; PRIOR FILING DATE: 2003-06-19

US-10-878-556A-75
 ; Sequence 75, Application US/10878556A
 ; Publication No. US20050286399A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hoffmann La-Roche Inc.
 ; TITLE OF INVENTION: HCV regulated protein expression
 ; FILE REFERENCE: 21762
 ; CURRENT APPLICATION NUMBER: US/10/878,556A
 ; CURRENT FILING DATE: 2004-06-28
 ; NUMBER OF SEQ ID NOS: 199
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 75
 ; LENGTH: 614
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: sw_hum/ddx5_human
 ; DATABASE ENTRY DATE: 1990-08-01
 US-10-878-556A-75

Query Match 3.2%; Score 207.5; DB 6; Length 614;
 Best Local Similarity 22.6%; Pred. No. 1.9e-06;
 Matches 110; Conservative 63; Mismatches 222; Indels 91; Gaps 19;
 Qy 387 GECFGGGATVTK-----ESCFLNQFDHAAQCPRASEEDTAVGPEPLVPS 436
 Db 31 GKKEGNGEKLKKNLDELKPEKNFYQHPD-----LARTAQEVETYSRSEKITVRG 86
 Qy 437 ---POPVEPVSLEDTVLYSLGSGQLAETPAEVFOALBQLGHQAFRPGQERAVMRIL 493
 Db 87 HNCPE-----VLFNFE-----ANFPANVMDVIARQNFTPTAIQAQGPVAL 129
 Qy 494 SGISTLLVPTGAGKSLCYQLPALLYRRSP-----CLTLVVSPLSLMDQVSG- 543
 Db 130 SGLDMVGVAQTGSGKTLSPALLPAIHHNQPFLEGRDGPICLVLAFTRELAQQVQVAAE 189
 Qy 544 -LPPC-LKAACTHSGMTRKQRESVLQIRAAQVHVLMTPEALVG--AGGLPPAAQLPP 598
 Db 190 YCRACRLKSTCIYGAPKGPQIRDLER---GVEICATPGRLIDFLECG---KTNLRR 241
 Qy 599 VAFACIDEAHCLSQSHNFRPCYLKRVCKVLRMGVHCFGLGTATATRTASDVAQ---H 655
 Db 242 TTYLVLDADRMLDM--GFEPQIRKIVDQIRPDRQTLKMSATWPKVEVQLAEDFLKDYIH 299
 Qy 656 LAVAEEDPLHGPAPVPTNLHL-----SVSMRDRTDQALLTLQKRFQNLDSIIICNRR 711
 Db 300 INI-----GALELSANHNLIQIVDVCHDVEKDEKLRLMEEIMSEKENKTIVEFV--E 349
 Qy 712 DTERIAALLRTCLHAAPVPGSGRAPKTTAEAYHAGMCSRERRRVRQAFMQGLRVVAT 771
 Db 350 TKRRCDLTRKWRDGM--PAMG-----IHGDKSQQERDQVNLNFKHGKAPILIAT 398
 Qy 772 VAFGMGLDRDPVRAVLHGLPPSFESYVQAVGRAGRDGQPAHCHLFLOP-----QGEDLRE 827
 Db 399 DVASGLDVEDVKFVINYDYPNSSEDIYHRIQRTARSTKTGTAYTFTFPNNIKQVSDLS 458
 Qy 828 LRRHVH 833
 Db 459 VLREAN 464

RESULT 14
 US-11-012-762-34
 ; Sequence 34, Application US/11012762
 ; Publication No. US20050244815A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Georgia State University Research Foundation, Inc.
 ; TITLE OF INVENTION: Compositions and Methods for Viral Resistance Genes
 ; FILE REFERENCE: GSU1.PCT
 ; CURRENT APPLICATION NUMBER: US/11/012,762
 ; CURRENT FILING DATE: 2004-12-15
 ; PRIOR APPLICATION NUMBER: PCT/US03/19300
 ; PRIOR FILING DATE: 2003-06-19

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:39:51 ; Search time 235.2 Seconds
(without alignments)
3623.632 Million cell updates/sec

Title: US-09-889-325-4

Perfect score: 6424

Sequence: 1 MERLRDVRERLQWERAFFR.....HLSFHALVGLATELLOVAR 1208

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05_80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------------|---------------------|
| 1 | 6424 | 100.0 | 1208 | 1 | RECQ4_HUMAN | O94761 homo sapien |
| 2 | 3944 | 61.4 | 1216 | 1 | RECQ4_MOUSE | Q75nr7 mus musculus |
| 3 | 2659.5 | 41.4 | 1500 | 2 | Q4JNX8_XENLA | Q4jnx8 xenopus lae |
| 4 | 2310.5 | 36.0 | 910 | 2 | Q4RLC3_TETNG | Q4rlc3 tetraodon n |
| 5 | 1723.5 | 26.8 | 1579 | 2 | Q9VSE6_DROME | Q9vse6 drosophila |
| 6 | 1680.5 | 26.2 | 1530 | 2 | Q9NH11_DROME | Q9nh11 drosophila |
| 7 | 1670.5 | 26.0 | 1058 | 2 | Q7PNN8_ANOGA | Q7pnn8 anopheles g |
| 8 | 1172 | 18.2 | 340 | 2 | Q7YR85_BOVIN | Q7yr85 bos taurus |
| 9 | 971.5 | 15.1 | 927 | 2 | Q7X829_ORYZA | Q7x829 oryza sativ |
| 10 | 897.5 | 14.0 | 870 | 2 | Q8W028_ARATH | Q8w028 arabidopsis |
| 11 | 872 | 13.6 | 941 | 2 | Q9C6N0_ARATH | Q9c6n0 arabidopsis |
| 12 | 761 | 11.8 | 874 | 2 | Q7FAH0_ORYZA | Q7fah0 oryza sativ |
| 13 | 635 | 9.9 | 704 | 2 | Q7MID7_VIBVU | Q7mid7 vibrio vuln |
| 14 | 631.5 | 9.8 | 608 | 1 | RECQ_ECOLI | F15043 escherichia |
| 15 | 631.5 | 9.8 | 611 | 2 | Q8X8N1_ECO57 | Q8x8n1 escherichia |
| 16 | 631.5 | 9.8 | 611 | 2 | Q8FBM6_ECOL6 | Q8fbm6 escherichia |
| 17 | 631.5 | 9.8 | 611 | 2 | Q83IW3_SHIFL | Q83iw3 shigella fl |
| 18 | 626 | 9.7 | 639 | 2 | Q8DBIC_VIBVU | Q8dbic vibrio vuln |
| 19 | 617.5 | 9.6 | 654 | 2 | Q7MBS5_VIBPA | Q7mbs5 vibrio para |
| 20 | 613.5 | 9.6 | 608 | 2 | Q6CZJ7_ERWCT | Q6czj7 erwinia car |
| 21 | 613 | 9.5 | 642 | 2 | Q87U96_PSESM | Q87u96 pseudomonas |
| 22 | 609.5 | 9.5 | 609 | 2 | Q823B2_SALTI | Q823b2 salmonella |
| 23 | 609.5 | 9.5 | 615 | 2 | Q57HP9_SALCH | Q57hp9 salmonella |
| 24 | 607.5 | 9.5 | 608 | 1 | RECQ_SALTY | P40724 salmonella |
| 25 | 607 | 9.4 | 644 | 2 | Q4KJZ8_PSEF5 | Q4kjjz pseudomonas |
| 26 | 604 | 9.4 | 649 | 2 | Q4ZLJ7_PSESY | Q4zlj7 pseudomonas |
| 27 | 604 | 9.4 | 654 | 2 | Q8EKL1_SHEON | Q8eek1 shewanella |
| 28 | 603.5 | 9.4 | 609 | 2 | Q5PKN6_SALPA | Q5pkn6 salmonella |
| 29 | 599 | 9.3 | 746 | 2 | Q7NHAB_GLOVI | Q7nhab gloeobacter |
| 30 | 594.5 | 9.3 | 595 | 2 | Q4NW16_9DELT | Q4nw16 anaeromyxob |
| 31 | 593 | 9.2 | 610 | 2 | Q66FY3_YERPS | Q66fy3 yersinia ps |

| | | | | | | |
|----|-------|-----|-----|---|--------------|--------------------|
| 32 | 593 | 9.2 | 610 | 2 | Q8ZAG8_YERPE | Q8zag8 yersinia pe |
| 33 | 592.5 | 9.2 | 464 | 2 | Q5DYC0_VIBF1 | Q5dyc0 vibrio fisc |
| 34 | 586.5 | 9.1 | 608 | 2 | Q7MYQ5_PHOLL | Q7myq5 photorhabdu |
| 35 | 586 | 9.1 | 602 | 2 | Q88Y35_LACPL | Q88y35 lactobacill |
| 36 | 584.5 | 9.1 | 641 | 2 | Q6LK00_PHOPL | Q6lk00 photobacter |
| 37 | 582.5 | 9.1 | 651 | 2 | Q5FNA4_GLUOX | Q5fna4 gluconobact |
| 38 | 581.5 | 9.1 | 620 | 2 | Q9KVF0_VIBCH | Q9kvf0 vibrio chol |
| 39 | 580.5 | 9.0 | 749 | 2 | Q7UR05_RHOBA | Q7ur05 rhodopirell |
| 40 | 578.5 | 9.0 | 632 | 1 | RECQ_PASMU | Q9cl21 pasteurella |
| 41 | 578 | 9.0 | 619 | 1 | RECQ_HAEIN | P71339 haemophilus |
| 42 | 577 | 9.0 | 619 | 2 | Q4QMG6_HAEI8 | Q4qmg6 haemophilus |
| 43 | 575 | 9.0 | 679 | 2 | Q5LWQ8_SILPO | Q5lww8 silicibacte |
| 44 | 573.5 | 8.9 | 615 | 2 | Q6LLP9_PHOPL | Q6llp9 photobacter |
| 45 | 573 | 8.9 | 637 | 2 | Q65QS2_MANSM | Q65qs2 mannheimia |

ALIGNMENTS

RESULT 1
RECQ4_HUMAN STANDARD; PRT; 1208 AA.
ID RECQ4_HUMAN STANDARD; PRT; 1208 AA.
AC O94761; Q96DW2; Q96F55;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE ATP-dependent DNA helicase Q4 (EC 3.6.1.-) (RecQ protein-like 4)
DE (RecQ4) (RTS).
GN Name=RECQL4; Synonyms=RECQ4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND TISSUE SPECIFICITY.
RC TISSUE=Testis;
RX MEDLINE=99097344; PubMed=9878247; DOI=10.1006/geno.1998.5595;
RA Kitao S., Ohsugi I., Ichikawa K., Goto M., Furuichi Y., Shimamoto A.;
RT "Cloning of two new human helicase genes of the RecQ family:
biological significance of multiple species in higher eukaryotes.";
RL Genomics 54:443-452(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], SUBCELLULAR LOCATION, AND DISEASE.
RX MEDLINE=20021764; PubMed=10552928; DOI=10.1006/geno.1999.5959;
RA Kitao S., Lindor N.M., Shiratori M., Furuichi Y., Shimamoto A.;
RT "Rothmund-Thomson syndrome responsible gene, RECQL4: genomic structure
and products.";
RL Genomics 61:268-276(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 468-1208.
TX TISSUE=Lymph, and Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., McQuellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[4]
 RN INDUCTION.
 RX PubMed=11032027; DOI=10.1038/sj.onc.1203841;
 RA Kawabe T., Teuyama N., Kitao S., Nishikawa K., Shimamoto A.,
 RA Shiratori M., Matsumoto T., Anno K., Sato T., Mitai Y., Seki M.,
 RA Enomoto T., Goto M., Ellis N.A., Ide T., Furuchi Y., Sugimoto M.;
 RT "Differential regulation of human RecQ family helicases in cell
 transformation and cell cycle.";
 RL Oncogene 19:4764-4772(2000).
 RN [5]
 RN DISEASE.
 RX PubMed=12952869; DOI=10.1093/hmg/ddz306;
 RA Siitonen H.A., Kopra O., Kaerlahti H., Haravuori H., Winter R.M.,
 RA Saanen A.-M., Peltonen L., Kestilä M.;
 RT "Molecular defect of RAPADILINO syndrome expands the phenotype
 spectrum of RECQL diseases.";
 RL Hum. Mol. Genet. 12:2837-2844(2003).
 RN [6]
 RN INTERACTIONS WITH UBR1 AND UBR2, SUBCELLULAR LOCATION, FUNCTION, AND
 RN IDENTIFICATION BY MASS SPECTROMETRY.
 RX PubMed=15317757; DOI=10.1093/hmg/ddb269;
 RA Yin J., Kwon Y.T., Varshavsky A., Wang W.;
 RT "RECQL4, mutated in the Rothmund-Thomson and RAPADILINO syndromes,
 RT interacts with ubiquitin ligases UBR1 and UBR2 of the N-end rule
 pathway.";
 RL Hum. Mol. Genet. 13:2421-2430(2004).
 RN [7]
 RN VARIANTS ASP-267 AND GLN-1005.
 RX PubMed=12601557; DOI=10.1007/s100380300016;
 RA Roversi G., Beghini A., Zambruno G., Paradisi M., Larizza L.;
 RT "Identification of two novel RECQL4 exonic SNPs and genomic
 RT characterization of the IVS12 minisatellite.";
 RL J. Hum. Genet. 48:107-109(2003).
 RN [8]
 RN VARIANTS GLN-355, SER-441 AND 857-CYS--THR-858 DEL.
 RX PubMed=15221963; DOI=10.1002/ijc.20269;
 RA Nishiyo K., Nakayama T., Aoyama T., Okamoto T., Ishibe T., Yasura K.,
 RA Shima Y., Shibata K.R., Teuboyama T., Nakamura T., Toguchida J.;
 RT "Mutation analysis of the RECQL4 gene in sporadic osteosarcomas";
 RL Int. J. Cancer 111:367-372(2004).
 CC -1- FUNCTION: DNA-dependent ATPase. May modulate chromosome
 CC segregation.
 CC -1- SUBUNIT: Interacts with UBR1 and UBR2.
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
 CC thymus and testis.
 CC -1- INDUCTION: Up-regulated in actively dividing cells.
 CC -1- DISEASE: Defects in RECQL4 are a cause of Rothmund-Thomson
 CC syndrome (RTS) [MIM:268400]. A disease characterized by
 CC dermatological features such as atrophy, pigmentation, and
 CC telangiectasia and frequently accompanied by juvenile cataract,
 CC saddle nose, congenital bone defects, disturbances of hair growth,
 CC and hypogonadism.
 CC -1- DISEASE: Defects in RECQL4 are a cause of RAPADILINO syndrome
 CC [MIM:266280]. A disease characterized by radial and patellar
 CC aplasia or hypoplasia.
 CC -1- SIMILARITY: Belongs to the helicase family, RecQ subfamily.
 CC -1- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Hamatol.;
 CC WWW="http://www.infobiogen.fr/services/chronocancer/genes/RECQL4ID285.html".
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC ENBL; AB006532; BAA74453.1; -; mRNA.
 CC ENBL; AB026546; BAA86899.1; -; Genomic_DNA.
 CC ENBL; BC011602; AAH11602.2; -; mRNA.
 CC ENBL; BC013277; AAH13277.2; ALT_INIT; mRNA.
 CC HSP; P15043; 10Y.
 CC Ensembl; ENSG00000160957; Homo sapiens.
 CC HGNC; HGNC:9949; RECQL4.

DR MIM; 603780; --
 DR MIM; 268400; --
 DR MIM; 266280; --
 DR GO; 0003678; P:DNA helicase activity; TAS.
 DR GO; 0007275; P:development; TAS.
 DR GO; 0006281; P:DNA repair; TAS.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR002464; DEAD/DEAH_box.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004589; RecQ.
 DR PANTHER; PTHR13710; RecQ; 1.
 DR Pfam; PF00270; DEAD; 1.
 DR TIGRFAMs; TIGR00614; recQ_fam; 1.
 DR PROSITE; PS00630; DEAD_ATP_HELICASE; FALSE_NEG.
 KW ATP-binding; Catabact; Helicase; Hydrolase; Nuclear protein;
 KW Nucleotide-binding; Polymorphism.
 FT NP BIND 502 509 ATP (Potential).
 FT MOTIF 605 608 DEAD box.
 FT VARIANT 267 267 E -> D (common polymorphism).
 FT VARIANT 355 355 R -> Q.
 FT VARIANT 441 441 P -> S.
 FT VARIANT 857 858 Missing.
 FT VARIANT 1005 1005 R -> Q (common polymorphism).
 FT FTID=VAR_023299.
 SQ SEQUENCE 1208 AA; 133077 MW; CB809A765A848A1 CRC64;
 Query Match 100.0%; Score 6424; DB 1; Length 1208;
 Best Local Similarity 100.0%; Pred. No. 2.7e-307;
 Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MERLDVRRERLQAWERAFRRQRPRSQDDVEAAPEETRALYREYRTLKRTTGQAGGLR 60
 DB 1 MERLDVRRERLQAWERAFRRQRPRSQDDVEAAPEETRALYREYRTLKRTTGQAGGLR 60
 QY 61 SSESLLPAAAEAEPEPCWGPFLNRAATKSPQTPGRSRQGSVPDYGQRLKANLKGTLQAG 120
 DB 61 SSESLLPAAAEAEPEPCWGPFLNRAATKSPQTPGRSRQGSVPDYGQRLKANLKGTLQAG 120
 QY 121 PALGRPWPPLGRASSKASTPKPGTGPVPSFAKVSDEPPQLPEPQPRGRLOHLOASLS 180
 DB 121 PALGRPWPPLGRASSKASTPKPGTGPVPSFAKVSDEPPQLPEPQPRGRLOHLOASLS 180
 QY 181 QRLGSLDPGWLQCHSEVPDFLGPACACRDLGSEESQLLIPGESAVLPGAGSQGPES 240
 DB 181 QRLGSLDPGWLQCHSEVPDFLGPACACRDLGSEESQLLIPGESAVLPGAGSQGPES 240
 QY 241 AFOEVSIRVSGPQSSSGGEGKRWNEEPWESPAQVQOESSQAGPPSEGAGAVAEEDPPG 300
 DB 241 AFOEVSIRVSGPQSSSGGEGKRWNEEPWESPAQVQOESSQAGPPSEGAGAVAEEDPPG 300
 QY 301 EPVOAOPPCSSPSNPRVHGLSPSSQARAGKAGCTAPLHIFPLARHHDGNYVRLNMKQ 360
 DB 301 EPVOAOPPCSSPSNPRVHGLSPSSQARAGKAGCTAPLHIFPLARHHDGNYVRLNMKQ 360
 QY 361 KHVYRGALRSRLLRQAWKQKWKCEGCGGGGATVTTKESCFLNEQFDHMAAQCPRPA 420
 DB 361 KHVYRGALRSRLLRQAWKQKWKCEGCGGGGATVTTKESCFLNEQFDHMAAQCPRPA 420
 QY 421 SEEDTDAVGPEPLVPSFPVPEVPSLDPTVLPYSLGPSQQLAETPAEVFQALEQLGHOA 480
 DB 421 SEEDTDAVGPEPLVPSFPVPEVPSLDPTVLPYSLGPSQQLAETPAEVFQALEQLGHOA 480
 QY 481 FRPQERAVWRILSGISTLLVLTGTAGKSLCYQLPALLYSRRSPCLTLVVSPLLSDMDQ 540
 DB 481 FRPQERAVWRILSGISTLLVLTGTAGKSLCYQLPALLYSRRSPCLTLVVSPLLSDMDQ 540
 QY 541 VSLGPLCLKAACIHSQMTKQRESVLQKIRAAQVHVLMLTPTPEALVGAGGLPPAAQLPPVA 600

Db 541 VSGLPPLKALACIISGMRKQBSVLQKIRAAQVHVLMLTPEALVAGGLPPAQLPPVA 600
Qy 601 FACIDEAHCLSQWHSNFPCLYRCKVLRRMGVHCFTGLTATATRTTASDAVQHLAAVE 660
Db 601 FACIDEAHCLSQWHSNFPCLYRCKVLRRMGVHCFTGLTATATRTTASDAVQHLAAVE 660
Qy 661 EPDLHGPAVPTNLHLSVSMRDPTDQALLTLLOKRFONLDSIIYCNRRDTERIAALL 720
Db 661 EPDLHGPAVPTNLHLSVSMRDPTDQALLTLLOKRFONLDSIIYCNRRDTERIAALL 720
Qy 721 RTCLHAAPVPSGGGRAPKTATZAYHAGCSRRRRVQAFMQGQLRVVVATVAFGMGLDR 780
Db 721 RTCLHAAPVPSGGGRAPKTATZAYHAGCSRRRRVQAFMQGQLRVVVATVAFGMGLDR 780
Qy 781 PDVRAVLHGLPPSPESYVQVAGRAGDQPAHCHLFLQPGEDLRELRRHVHADSTDFL 840
Db 781 PDVRAVLHGLPPSPESYVQVAGRAGDQPAHCHLFLQPGEDLRELRRHVHADSTDFL 840
Qy 841 AVKRLVQRFVFACTCTCTRPSEQEGAVGGERPVKYPQAEQIASHQAAPGPRVCMGH 900
Db 841 AVKRLVQRFVFACTCTCTRPSEQEGAVGGERPVKYPQAEQIASHQAAPGPRVCMGH 900
Qy 901 ERALPIQLTVQALDMPERAIETLLCYLPHHMLLELLATTYTHCLMCPGPAQLQALA 960
Db 901 ERALPIQLTVQALDMPERAIETLLCYLPHHMLLELLATTYTHCLMCPGPAQLQALA 960
Qy 961 HRCPPPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWHDHEPTGV 1020
Db 961 HRCPPPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWHDHEPTGV 1020
Qy 1021 RRGTVLVEFSELAFLHSPGDLTAEKDKQICDFLYGRVQARERQALRLRTFOAFHSV 1080
Db 1021 RRGTVLVEFSELAFLHSPGDLTAEKDKQICDFLYGRVQARERQALRLRTFOAFHSV 1080
Qy 1081 APPSGCPLCQDDRERTRLDLGRYFEEREGSPGGMEDAQGPQCARLQWEDQVR 1140
Db 1081 APPSGCPLCQDDRERTRLDLGRYFEEREGSPGGMEDAQGPQCARLQWEDQVR 1140
Qy 1141 CDIROFLSLRPEEKFSRAVARIFHGIGSPCPYPAQVYQDQRRFWKYLHLSFHALVGLAT 1200
Db 1141 CDIROFLSLRPEEKFSRAVARIFHGIGSPCPYPAQVYQDQRRFWKYLHLSFHALVGLAT 1200
Qy 1201 BELLQVAR 1208
Db 1201 BELLQVAR 1208

RESULT 2
RECQ4 MOUSE
ID RECQ4 MOUSE STANDARD; PRT: 1216 AA.
AC Q75NR7; Q76MT1; Q99PV9;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE ATP-dependent DNA helicase Q4 (EC 3.6.1.-) (RecQ protein-like 4).
GN Name=Recq4; Synonyms=Recq4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA] (ISOFORM 1) AND NUCLEOTIDE
RP SEQUENCE [MRNA] OF 16-1115 (ISOFORM 2).
RX PubMed=11167012;
RA Ohnata T., Araki R., Fukumura R., Kuroiwa A., Matsuda Y., Tatsumi K.,
RA Abe M.;
RT "Cloning, genomic structure and chromosomal localization of the gene
RL encoding mouse DNA helicase RecQ protein-like 4.";
RN Gene 261:251-258(2000).
RP FUNCTION.

RX PubMed=12915449; DOI=10.1093/hmg/ddg254;
RA Hoki Y., Araki R., Fujimori A., Ohnata T., Koseki H., Fukumura R.,
RA Nakamura M., Takahashi H., Noda Y., Kito S., Abe M.;
RT "Growth retardation and skin abnormalities of the Recq4-deficient
RL mouse.";
RL Hum. Mol. Genet. 12:2293-2299 (2003).
RN [3]
RP DEVELOPMENTAL STAGE.
RX PubMed=12952869; DOI=10.1093/hmg/ddg306;
RA Siitonen H.A., Kopra O., Kaeserlaenen H., Haravuori H., Winter R.M.,
RA Saenanaen A.-M., Peltonen L., Kestila M.;
RT "Molecular defect of RAPADILINO syndrome expands the phenotype
RL spectrum of RECQL diseases.";
RL Hum. Mol. Genet. 12:2837-2844 (2003).
RN [4]
RP FUNCTION.
RX PubMed=15703196; DOI=10.1093/hmg/ddi075;
RA Mann M.B., Hodges C.A., Barnes E., Vogel H., Hassold T.J., Luo G.;
RT "Defective sister-chromatid cohesion, aneuploidy and cancer
RL predisposition in a mouse model of type II Rothmund-Thomson
RT syndrome.";
RL Hum. Mol. Genet. 14:813-825 (2005).
CC -!- FUNCTION: DNA-dependent ATPase (By similarity). May play a role in
development of the palate and the limbs. May modulate chromosome
segregation.
CC -!- SUBUNIT: Interacts with UBR1 and UBR2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q75NR7-1; Sequence=Displayed;
Name=2;
IsoId=Q75NR7-2; Sequence=VSP_015177;
CC -!- DEVELOPMENTAL STAGE: Not expressed at E12.5. Expressed at E15.5-
E18.5, with highest levels in chondrocytes of developing bone and
cartilage and immature proliferating enterocytes of intestine.
CC -!- MUSCLE-SPECIFIC: Absence of the protein causes early embryonic
lethality. Transgenic mice with exon 13-deleted RECQL4 are severely
growth-retarded and show high (95%) perinatal lethality. They
exhibit various skin, bone, intestine, tooth and thymus
abnormalities and premature aging features, but have normal
sensitivity to IR and UV irradiation. In contrast, transgenic mice
expressing a truncated form of RECQL4 exhibit mild perinatal
lethality, no growth defect, but show defects of the skin and
skeleton, aneuploidy and increased cancer susceptibility.
CC -!- SIMILARITY: Belongs to the helicase family. RecQ subfamily.
CC -!- SIMILARITY: Contains 1 CCHC-type zinc finger.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; AB039882; BAD11131.1; -; mRNA.
DR EMBL; AB175741; BAD14289.1; -; mRNA.
DR EMBL; AB042529; BAB32696.1; -; Genomic_DNA.
DR HSSP; P15043; 10Y7.
DR Ensemble; ENSMUSG0000033762; Mus musculus.
DR MGI; MGI:1931028; Recq4.
DR PROSITE; PS00690; DEAH_ATP_HELICASE; FALSE_NEG.
DR PROSITE; PS50158; ZF_CCHC; 1.
KW Alternative splicing; ATP-binding; Helicase; Hydrolase; Metal-binding;
Nuclear protein; Nucleotide-binding; Zinc; Zinc-finger.
FT ZN FING 393 410 CCHC-type.
FT NP BIND 519 526 ATP (Potential).
FT MOTIF 627 630 DEAH box.
FT VARSPLIC 843 843
FT CONFLICT 732 739
FT CONFLICT 878 878 S -> G (in Ref. 1; BAD11131).


```
Db 1419 VPGDFTSEIDNALDTLTYTSVKQERTQLQLQYVAHGLAAVAYSSCGCCNADFPQDRG 1478
Qy 1098 TRKXDLGRYFEEBEGQ-----EPGWMEDAQGPQOARLQDWEDOVRCDIRQFLSLRP 1151
Db 1479 EQLKAIVRNYFANDYPQDLELEIPESNVDP-----ENIIDVHALINNYP 1523
Qy 1152 EEPSSRAVARIFHGICSPCAQVQDQRRFRWKYHLHLSFHALVGLATEBELLQ 1205
Db 1524 DNTFTGRNIARIPIHGIMSPNYPVWGR-CKFWRAHVKVDFNRIHLANMAIIK 1576

RESULT 6
Q9NH11 DROME
ID Q9NH11 DROME PRELIMINARY; PRT; 1530 AA.
AC Q9NH11;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RECQ4.
GN Name=RecQ4; Synonyms=RECQ4; ORFNames=CG7487;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kueano K., Berres M.E., Engels W.R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF233026; AA42939.1; -; Genomic_DNA.
DR HSSP; P15043; 10Y.
DR IntAct; Q9NH11; -.
DR FlyBase; FBgn0040290; CG7487.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00098; zf-CCHC_1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR TIGRFAMs; TIGR00614; recQ_fam; 1.
DR PROSITE; PS0158; ZF_CCHC_1.
SQ SEQUENCE 1530 AA; 169878 MW; 1B753A59A3E30EB4 CRC64;

Query Match 26.2%; Score 1680.5; DB 2; Length 1530;
Best Local Similarity 32.4%; Pred. No. 3.6e-74;
Matches 420; Conservative 191; Mismatches 412; Indels 275; Gaps 36;

Qy 107 QRLKANIUKTGLQAGP-ALGRRPWPLGRASSKASTPKPGTGPVPSFAEKVSDREPPQLPEP 165
Db 306 QELQTDNNSSNNQKPDHLNHTP-PASSQKSVAPKNKPP-----PSEQETSDSDSVVAES 359
Qy 166 QPRGRQLQHLQASLSQRLGSLDQWLQ-RCHSEVPDFLGAPKAC-----RPLDGEESQ-- 218
Db 360 EEEPEPEYQLSKRKRIIVTASKEVVAAPVEIPNKEVETETETFAQENDFSADEQDA 419
Qy 219 LLIP-----GESAVLPGAGSQ---GPESAFAPQEVSVIRVSGPQPSGSGGKKRWNEEPWES 271
Db 420 TYVPENKKDKAKRKAQAGKQKTKPKA-----EPKPKT-----EK 455
Qy 272 PAQVQOESSQAG--PPSEGAGAVAEEDPGEFPVQAQPPQP-----CSSPSNDR 318
```

```
Db 456 KAKVKAEEKPKAEKKPRNSKKAIAVEPAPDPEDERQPLNPEDLKVYLALAEAGDITSVPR 515
Qy 319 YH-----GLSPSSQARAGKAGTAPLHIFPPLARHD-RGNY 353
Db 516 INVQDLEEDADATAQRIYRTFAAGPNFGLSEGSNIRVDEKKAARAAKKLEERIAAGKLNENF 575
Qy 354 VRLNMKQKHVVRG-RALRSRLLEKQAWKQWR-----KKEGCEFGGGATVTTKESC 403
Db 576 VTINIQKKFVRGKSKVNFSSKYKKQWRHKRVAALSGLPMDMGCGDGG---VLT---C 628
Qy 404 FLNEQDFHWAQC-----PRP----- 419
Db 629 FQCGGVGHGFAQCKVKGDSLLPLSAQLLEDPSFPPTLAEAEQEMASQGVAVHRSNIRSLP 688
Qy 420 -----ASEEDTDAVGPEPLVPSQP----- 439
Db 689 QAANAAILQDELNESEEDDOESSEEEVQOHPDPNWSDEMVDVDFALDAAVEASLSQP 748
Qy 440 -----VPE-----VPSLDPT-----VLPLYSLGSPGQLAETPA 467
Db 749 VSQEAASPIKTYVGHKIPEEFPLKQAGLDTTASSNSRSHQGGVKPLVDLLPDGVSQDTPP 808
Qy 468 EVFQALQLGHOQAFRPGQERAVNRILSGISTLLVLTGACKSLCYQLPALLYRRSRCLT 527
Db 809 EVLEALHMFHTNFRKGQDRAIMRTLSGLSSVLTSTGSGKSLCYQLPAYLYSRKVGAIT 868
Qy 528 LVVSPILSLMDDQVGLPPCLKAACIHSQWTRKQRESVLQKIRAAQVHVLMLTPEALV-- 585
Db 869 LVISPLVSLMEDQVGVPHFLRAHCLHTNQTAQRMKIQMIANGEIDILLVSPFAVVAG 928
Qy 586 ---GAGGLPPAQLPPVAFACIDEAHLCSQWGHNFRPCYLRCVKLRRMGRVHCFGLGT 641
Db 929 ERATGFGAI--LRQLPPIAFACIDEAHCVSQWSHNFRPSYLMICKVLKRLKGLVRLTGLT 986
Qy 642 ATATRTASDVAOHLAVB--EPDLHGPAVPPTNLHLHSVMDRDTQDQALLTLLOGKRQNL 700
Db 987 ATATLPTRVSIINHLGISDGERGIISDIPLDNVLVSVSKDENRDALLQLLNSERPEPC 1046
Qy 701 DSIIICNNRREDTERIAALLRTCLCHAAWVPGSGRAPKT---TAAAYHAGMCSRERRV 756
Db 1047 QSIIICVTRDECEIRAGFIRTCVQDRREPTQOTKGRKRVNQAEPIYHAGMFSRRRTV 1106
Qy 757 QRAFMQQLRVVAVTAVFGMLDRDPVRAVLHLGLPSPFESYVQAVGRAGRDGQPARCHL 816
Db 1107 QKAFMSNELRIVATIAFGMGINKPDIRAVIHYNMPNPFESYVQEIGRAGRDGLPSHCHL 1166
Qy 817 FLOPQGEDLRELRRHVHADSTDFLAVKRLVQRVPPACTCTCTPPPEQEGAVGGERVPVK 876
Db 1167 FLDAKGQDQSELRRHVYSNIDRHVIRKLIQKIFVPCSC-----DKEASKRTALP- 1218
Qy 877 YPPQEAQLSHQAAPGPR-RVCMGHERALPIQLTVQALDMPHEAETLLCYLELHPHHL 935
Db 1219 -----LEGDGRVIMCGHIGSVKTEVEMLDIPAENISTLLCYMELDPRWC 1267
Qy 936 ELIATTYTHCLNCPGCPAQALQALAHRCPPAVLAQLQALPDP--GQSSSVFEDVMKLVD 994
Db 1268 SVLSSAVMAKVISYGGPKYLKHAACECPPLAMAIQIIRDKTFKEDSNIIIESVTDIAA 1327
Qy 995 SMGWELASVRRALCOLQWDHEPRTGVRRTGVLVEFSELAFLHRSQGLTAEKQDQICDF 1054
Db 1328 GIGWNSGVVYQLKOLEW--VKYNGYPKRSPITVSFVLDGFRIKVPGDFTSEIDNALDT 1385
Qy 1055 LYGRVQARERQALARRTRTFQAFHSVAFPSGCPGLEQ--DEERSTRKDLILGRVFESEEG 1113
Db 1386 LYTRSVKQERTQLQIYVAHGLAAVAYSSCGCCNADFPQDRGEQLKAIVRNYFANDYP 1445
Qy 1114 Q-----EPGMBDAQGPQARLQDWEDQVRCDIRQFLSRPEEKFSRAVARIFHG 1167
Db 1446 QDLELEIPESNVDP-----ENIIDVHALINNYPDNTFTGRNIARIPIHG 1490
Qy 1168 GSPCYPAQVYQDRFRWRKYLHLSFHALVGLATEELLQ 1205
Db 1491 MSPNYPVWGR-CKFWRAHVKVDFNRIHLANMAIIK 1527
```

```

RESULT 7
Q7PNM8 ANOQA PRELIMINARY; PRT; 1058 AA.
AC Q7PNM8;
DT 01-MAR-2004 (TremBLrel. 26, Created)
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE ENSANGP0000004546 (Fragment)
GN ORFNames=ENSANGG00000003572;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RL CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AA0801008960; EAA11913.3; --; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00270; DEAD; 1. CCHC.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR TIGRfam; TIGR00614; recQ_fam; 1.
DR PROSITE; PS0158; ZF_CCHC; 1.
KW ATP-binding; Helicase; Hydrolase.
FT NON TER 1
SQ SEQUENCE 1058 AA; 117711 MW; 6DFBPF17EA2255FC CRC64;

Query Match 26.0%; Score 1670.5; DB 2; Length 1058;
Best Local Similarity 37.5%; Pred.No. 7.4e-74;
Matches 386; Conservative 147; Mismatches 304; Indels 191; Gaps 24;

QY 321 GLSPSSQARAGAEAGTAPLHIFPRLARHD-RGNYVRLNMKQHYVRG-RALRSLLRLKQA 378
DB 78 GPSSSGKTKTTVAGSTKETALRKKMAAGKLNENFRIDIRKKVFKVKKTKTINYRYKKSQ 137
QY 379 WKQKRWKGEFCFG-----GGATVTVTKSCFLNEQPDHWAACPRPABEE---DTD 426
DB 138 WK---AKKAAALTGPMDNRGCDGGTLT---CFQCGGTGHMAFCKPLEDKLLPYDAD 190
QY 427 AVGPE-----PLVPSP-----QPVEF 442
DB 191 TIESSFPTEAEAMANSRTLAVHSNRTERLPVANPVWKEQEQBDLPABEEVGEPGKE 250
QY 443 VPSLDP-----448
DB 251 EPSTLPVEDDLDEPMEQSDHQTEQNAASQAPAPAYIGHKIPPEFLKQSGLLVTV 310

```

```

QY 449 -----TVLPYLSLPGSQLAETPAEVFOALEQHQAPRGOERAYVMRILSGISTILLVL 502
DB 311 SGGFKGTVEPLYQORGDCTVPTTPEVFEALRMFGHOSFRHGQERAVNRVLCGLUSTITVL 370
QY 503 PTGAGKSLCYQLPALLYSRSPCLTLVVSPLLSLMDQVSGLPCLKAACIHSQMTKRQR 562
DB 371 STGAGKSLCYQLPAYLYRKNRSCITLVIPLVSLMEDQVHGMPDFLNAHCLHTTQTPKVR 430
QY 563 ESVLQKTRAAQVHVMULTPEALV-----GAGGLPPAAQLPPVAFACIDEAHCLISQSHN 616
DB 431 DRTMQAIAAGEVDVLLISPEAVSSEKSTGFGSL--LRQLPPIAFACIDEAHCVSQSHN 488
QY 617 FRPCYLAVCKVLRERMGVHCFGLGTATATRTASDVQAHLAVAEEDP-----LHGPAPVPT 672
DB 489 FRPSYLMICKVLKEKGVKLTILGTATATVQTRQISVSHLAI---PDGVQGIISDIPDP 545
QY 673 NLHLSVSMRDRTDQALLTLQGRFQNLDSIIICYNRREDTERIAALLRTHCLHAAPVGS 732
DB 546 NLLLTVSERDANRDVALVELLQSERFSSLSQSIIVYCTRRDDCERVATFTRTCFQDAARAA 605
QY 733 GGRAPKT-----TAEAYHAGCRRRRRVRQAFMQGLRVVAVAFQMGMLDRDPDRAVL 787
DB 606 AANAHKRKLNYVAEPYHAGMPASRRRTIQNAFMSGELRIVVATIAFGMGINKADIRAI 665
QY 788 HLGLPSPFESVVOAVGRAGRDGPQAHCHLFLQPOGEDLRELRHHVHADSTDFLAVKRLVQ 847
DB 666 HYNMPPKPFESVVOAVGRAGRDGLLHCHLFLQPOGEDLRELRHHVHADSTDFLAVKRLVQ 725
QY 848 RVFPACTCTCTRPPEQ-----EGAVGGERPVPKYPQEAQSLHQAAPGPRVCMGHE 901
DB 726 KIFVPCACA-----KSHQOVILFTDGGGGGT-----KRLCPGHE 760
QY 902 RALPIQITVQALDMPERAEITLLCYLELHPHHWLELLATTTHRLNCPGPGPAQLQALAH 961
DB 761 ICFSEIATVQQLDIPENITTFCLYELDEQRYIQAISPAYTMCKVMSYGVVPLRQAAK 820
QY 962 RCPPLAVCLAQOLPEDPGQG---SSVPEFDMVKLVDSMGWELASVRRALCOLQMDHEPTGV 1020
DB 821 ECPPLAMAFALDLKRGISHATSTATIEFPVIDVASAIGMDSGVVKQLKNLEW---TTVNV 878
QY 1021 RRGTVLVEFSELAFLHRS PGDLTAEEKDQICDFLYGRVQARERQALRLRTTFOAFHSV 1080
DB 879 RKRSPLSVTFTFELGFRVAPGDLTDELDHALDGLYERVTHQERTQLAQLOVISDALNSV 938
QY 1081 APPSGCPCLEQD---EERSTRLKOLLGRYFEEEGQEPGMEDAQPGPGQARLODWDQ 1138
DB 939 CFNTIGP-VSRADCPGSDKLTIVREYFTTDSKE---QIEIVPEPDDTT-----DEQ 989
QY 1139 VRCDIRQFLSLRPEKPSRAVARIFHGIGSPCYPAQVYQDRRFRWKYHLHSFHALVCL 1198
DB 990 LINDIRTAICRYPENFTGTRAIARLPHGVQSPNYTALVMSRS-NFWRAYTTKTFDNRI 1048
QY 1199 ATEELLQV 1206
DB 1049 ANAEIVRM 1056

RESULT 8
Q7YR85 BOVIN PRELIMINARY; PRT; 340 AA.
AC Q7YR85;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE RecQ protein-like 4 (Fragment).
GN Name=RECQL4;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]

```

RP NUCLEOTIDE SEQUENCE.
RX PubMed=14667821; DOI=10.1016/S0898-7543(03)00238-6;
RA Winter A., Alzinger A., Fries R.;
RT "Assessment of the gene content of the chromosomal regions flanking
RL bovine DCAT1";
RL Genomics 83:172-180(2004).
DR EMBL; AJ518973; CAD58810.1; -; Genomic_DNA.
DR HSP; P15043; 10YV.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00490; HELIC_C; 1.
FT NON TER 1
FT NON TER 340 340
SQ SEQUENCE 340 AA; 37366 MW; 205F057DD109CAED CRC64;
Query Match 18.2%; Score 1172; DB 2; Length 340;
Best Local Similarity 67.8%; Pred. No. 6.6e-50;
Matches 232; Conservative 27; Mismatches 57; Indels 26; Gaps 2;
Qy 601 PACIDEAHCLSQSHNFRPCVLRVCKV-----LRRMGVHC 636
Db 1 FACIDEAHCLSQSHNFRPCVLRVQVSLGGVSGVQREGTAGSPPPSOTLRDQMGVHC 60
Qy 637 FLGTATATRTASDVQAHLAVAEEDPLHGPAPVPTNHLVSVSDRDTQALLTLLOKGR 696
Db 61 FLGTATATRTASDVQAHLAVAEEDPLHGPAPVPTNHLVSVSDRDTQALLTLLOKGR 120
Qy 697 FQNLDSIIYCNRRDTERIAALRTCLHAAVPGSGRAPKTTAEAYHAGMCSRERRV 756
Db 121 FRALGSVIIYCHRRDTERVAALRTCLDAQDPGPHGRALEAVAEYHAGMCSRERRV 180
Qy 757 QRAFMQGLRVVAVTAFVGMGLDRPDVRAVLHGLPSPFESYVQAVGRAGRDGQPAHCHL 816
Db 181 QRAFMQGLRVVAVTAFVGMGLDRPDVRAVLHGLPSPFESYVQAVGRAGRDGQPAHCHL 240
Qy 817 FLOPQGDRLRLRHVHADSTDFLAVKRLVQRVEPACTCTCTPPSQEAGVGRVVPK 876
Db 241 FLRPOGDLRRHRHVHADVDFFAVRLVQRVEP--PCACARQPPQEGSRSEEGHLAG 298
Qy 877 YPOEAEQLSHQAAPGPRRVCMGHERALPIQLTVQALDMPPE 918
Db 299 APVAASQDSQSPVHTPCPGHERVLPVQTVQALDMPPE 340
RESULT 9
ID Q7X829 ORYSA PRELIMINARY; PRT; 927 AA.
AC Q7X829;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE QJ991113 30.2 protein.
GN Name=Ory991113 30.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=2237377; PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RL Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Wang Q., Zhang L., Lu Y., Mu Y., Zhang L.S., Yu Z., Fan D.,
RL Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Yu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,

RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4";
RL Nature 420:316-320(2002).
CC -|- FUNCTION: E1F4A is both a subunit of a high molecular weight
CC protein complex involved in cap recognition and is required as a
CC single polypeptide chain for mRNA binding to ribosome. It is an
CC ATP-dependent single stranded DNA-binding protein with a sequence-
CC independent unwinding activity (helicase) (By similarity).
CC -|- SUBUNIT: e1F4P is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least E1F4A, E1F4B and E1F4G (By similarity).
CC EMBL; AL662946; CAD41320.2; -; Genomic_DNA.
DR HSP; P15043; 10YV.
DR Gramene; O7X829; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC_C; 1.
DR TIGRFAMs; TIGR00614; recQ_fam; 1.
DR ATP-binding; Helicase; Hydrolase; Initiation factor;
DR Protein biosynthesis.
SQ SEQUENCE 927 AA; 102031 MW; 48B7BC5504C04FFA CRC64;
Query Match 15.1%; Score 971.5; DB 2; Length 927;
Best Local Similarity 26.5%; Pred. No. 1.4e-39;
Matches 303; Conservative 168; Mismatches 392; Indels 279; Gaps 37;
Qy 110 KANLKTQAGPALGRRPWPGLGRASSKASTPKPP---GTGVPSPFAEKVDEPPQLPEPQ 166
Db 13 EGSLLSDVSASPPRRSP---PRPAPPPPPPPPKHTRPAAATPKPKLPTTAAASAPA 69
Qy 167 PRPGRQLHQASLSQRLGSLDP--GWLQRCHSEVPDFLGAPKACRPDLGSEES---QLLIP 222
Db 70 PAPPPPTLRAAALS-----DPHGLAAR-----IAAGSALTAASGTASSSFRLLVQ 116
Qy 223 GESAVLPGAGSQGPASAFQVYSIRVSGSPQSSSGEKERRMNEEPWESPAQVQESSQA 282
Db 117 SRNPSFDPATFTAPASSAPSEVPSAAPRPPTAA-----TDAPPQTRPKRVHP 165
Qy 283 GPPSEGAGAVAEEDPPGPEVQAPQPPQPCSPSPNRYHGLSPSSQAPAGKAETAPLHIF 342
Db 166 NSVSEVAAAASAAAEQP-----KRAAGG----- 187
Qy 343 PRLARHDRGNVRLNMKQKHVYVGRALRSLRLKQAKWKKGECFGGGATVTTKES 402
Db 188 -----SEGNFVRLNI--NGYGRRTFKNSQAKRSTKCSWRRO-----RAAGAT----- 229
Qy 403 CFLNEQFDHWAACQPRPASEEDTDVGPPEPLVSPQVPEVPSLDPTVPLYSGLSPGSQL 462
Db 230 -----PRSQGDEGDLVAELLEREKQAASD-----SVLEAVE-----SV 264
Qy 463 AETPAEVFOALEOL-----GHQAFRFGQERAVNRLISGISTLLVLPFGAKSLCY-OLPA 516
Db 265 REDPSE--QNLKSLNAAAYGHDSPROGLEAIQOIVAGESTMLVLPFGAKSLCYQVPA 322
Qy 517 LLYRSRSPCLTVVSPILLSIMDQVSLPCLAKAACIHSQWTRKQRESVLQKTRAAOVHV 576
Db 323 MIL-----PGLTVVSPILLSIMDQVSLPCLAKAACIHSQWTRKQRESVLQKTRAAOVHV 378
Qy 577 LMLTPEALVGAGGLPPAAQLPVFAFACIDEAHCLSQSHNFRPCVLRV-CKVLRMRMGVH 635

Db 379 LFVSPERPLNEFEFLIIRDTLPISLVAIDEAHCISEWSHFRPSYLRASLRKLNQV 438
Qy 636 CFLGLTATARTASDVQAHLAVAEEDLHGFAPVPTNLHLSDMRDITDQALLTLQKG 695
Db 439 CILAWTATATQTLEIMWALEIFSD-NLIQTSQIRENLQLSISTDNRLKDLMLLKSP 497
Qy 696 RFQNLDSIIICNRREDTERIAALLTCLHAAWVPGSGRPAKTTAAAHGMCRRRR 755
Db 498 PFVDMRSIIIVCKFQAETDFVSKYL--CDN-----NITAKSYHSLGIKRSR 543
Qy 756 VQRAFQCOLRVVATVAFGMLDPRDRAVLHLGLPPSPESYVQAVGRAGRDGPAHCH 815
Db 544 VOELFCNSKIRVVATVAFGMLDSDVEGVHISLPSLEBEYIQETGRAGRDGLSHCH 603
Qy 816 LFLQFQGEDLELRHRRHADSTFLAVKRLVQRFPPA-----CTCTCTRPPEGEVAGG 870
Db 604 LLL--DSATFYKIRSLSHSDGVGAMSKFLYQIFSSENTTGCICSLAK----- 650
Qy 871 ERVPKYPQBAEQLSHQAPGPRVCMGHERALPIQLTVQALDMPHRAIETLLCYLELH 930
Db 651 VOELFCNSKIRVVATVAFGMLDSDVEGVHISLPSLEBEYIQETGRAGRDGLSHCH 674
Qy 931 PHWLELLATTYTHCLNCPGPAQLALAHRCPLAVC-----LAQQLPEDPGQSSSVE 986
Db 675 DQQIRLLPQPSVTCITL-----YFKTSPQLLADKDIILRSVLRSEMKDGHYV 723
Qy 987 FDMVKLVDSMGWELASVRRALCOLQMDHEPTGVRRGTGVLFVFS-ELAFHLRSPG--- 1041
Db 724 FDIPIRIANDLKITWNEV-----FDHLHK-----LKFSGEISFELKDPAYCYV 765
Qy 1042 -----DLTAEKDOI CDLYGRVQARERQALRLRTTQAFHSAVPSC----- 1085
Db 766 ILWRPDDFNALSAN-----LTKWLSEVSSIKSLDAMF-ALANFAVKGRKRTGCGSGQ 819
Qy 1086 -GPCLEQODEERSTRLLDGLGRYFEEBEGQEPGGMEDAQGPQARLDWEDOVRCDIR 1144
Db 820 HTPCIQKIME-----YFSKDDGTSENC-----RTQLQKSSPFIQADIK 859
Qy 1145 QFLSLRPEKFSRAVARIFHGIGSPCYPAQVYGQDRFRWFKYHLHLSFHALVGLATELL 1204
Db 860 VPIQSNFAKTPTRAVARIMHGISSPAPSVTWSKN-HFWGRYVEVDPLVWEAKAELV 918
Qy 1205 QV 1206
Db 919 KL 920

RESULT 10
Q8W028 ARATH PRELIMINARY; PRT; 870 AA.
AC Q8W028;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Helicase.
GN Name=rdl-5;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hartung F., Pichova H., Puchta H.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Eif4a is both a subunit of a high molecular weight
CC protein complex involved in cap recognition and is required as a
CC single polypeptide chain for mRNA binding to ribosome. It is an
CC ATP-dependent single stranded DNA-binding protein with a sequence-
CC independent unwinding activity (helicase) (By similarity).
CC -!- SUBUNIT: eIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.

CC It is composed of at least EIF4A, EIF4E and EIF4G (By similarity).
DR EMBL; AJ421618; CAD13472.1; -, mRNA.
DR HSSP; P15043; IOY.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002464; DEAH_box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR TIGRFAMs; TIGR00614; recQ_fam; 1.
DR PROSITE; PS00690; DEAD ATP HELICASE; UNKNOWN_1.
KW ATP-binding; Helicase; Hydrolase; Initiation factor; Nuclear protein;
KW Protein biosynthesis.
SQ SEQUENCE 870 AA; 96383 MW; 8B5ADA4F47A81723 CRC64;

Query Match 14.0%; Score 897.5; DB 2; Length 870;
Best Local Similarity 26.7%; Pred. No. 5.6e-36;
Matches 295; Conservative 160; Mismatches 368; Indels 281; Gaps 38;
Qy 133 ASSKASTPKPGTGPVPSFAEKVSDPEPQLPEPQPRGRQLQHLQASLSQRLSDLPGLQ 192
Db 3 SSSRSRKPAP-THPPNPSPQEAQVPSPPPPSP-----LFTNLPFRICQSPA--- 53
Qy 193 RCHSEVDFLGAFCACPDLCSEESQLLIPGESAVLPGAGSQSPESAFAFEVSRVGP 252
Db 54 RFSVSSVSP---SRLCSRASFTSVKLSGDGVFV-----PEPLVEVIA----- 95
Qy 253 QPSSSGGKKRWNEBPWESPAQVQESSQAGPSEGAGAVAEEDPPG---EPVQAQPPQ 309
Db 96 -----PPK-----SVRRKPNLITDTTSPVK 118
Qy 310 PCSSPSNPRVHGLSPSSQARAGKAGTAPLHIFRLARHGRGNVRLNMKKHYVRGAL 369
Db 119 PMWFRSN-----GNCE-----GNFVKLNNGK---RGKFF 145
Qy 370 RSRL-----LRQAWKQKWKKGECFGGGATVTTKESCFNLQFDHWAACQPRPASEE 423
Db 146 PSKYGVSKRSYSFRKRYKKKADGDGSL--EESDLQKI-----ED 191
Qy 424 DTDVAGPEPLVSPQVPEVPSLDPTVLFLYSLGSPSQLAETPAEVFOALSQGHQAFRP 483
Db 192 EANGF-----ISSVEDAILAVKTEASDENLTKLLNVY-----GYDSFRD 231
Qy 484 GQERAVNRILSGISTILVLPFGACKSLCYQLPALLYRSRSPCLTIVYSPILSLMDQVSG 543
Db 232 GLOQAIKMLGGSTMLVLPFGACKSLCYQIPAMIL-----EGITLWVSPILSLMDQKH 287
Qy 544 LPPCLKAACHTSGMTRKQRESVLQIRAAOVHVLMLTPEALVGAGGLPPAAQLPVPVAFAC 603
Db 288 LPSIIKGLISSSQRPPEATEFLRKEGIIKVLVSPERLLNVFELSMFMSLSVSLV 347
Qy 604 IDEAHCLSQWSHNRPFCYLRV-CKVLRERMGVHCFGLGTATATRTTASDVAAQLHAVASEP 662
Db 348 VDEAHCYSEWSHNFRPSYMLKASMLFSELKACILAMTATATMTLQAVMSLEI-PST 406
Qy 663 DLHGAPVPTNLHLVSMVSM---DRDTQALLTLQKRFQNLDSIIICNRREDTERIAAL 719
Db 407 NLIQKSQLRDNFELVSLSGANRMKD---LLTIMESPPYKEIRSIIVYCKFYQFQETDMSKY 464
Qy 720 LRTCLHAAWVPGSGRAPKTTAAEAHAGMCRRRRVORAFMQGLRVVAVTAFGMGLD 779
Db 465 LRD-----NNINAKGVHSLGPKADRVRIQESFCSNKRIRVVAVTAFGMGLD 510

QY 901 ERALPIQUTVQALDMPERAIETILCYLELHPHHLELLIATYTHCRNLCPGPAQLOALA 960
 Db 651 --SLVIESASQKFMKEVMQITLTHLGEVQVLRMLPQLNICOTLN-----F 697
 QY 961 HR--CPLAVLAQQLPEDPGQSSSVFDPVVKLVDSMG---WELASVRALCOLQWDH 1014
 Db 698 HKTSDAQINVMYVQSSNTLAARSALVAAILKSHVKQGLHVFDPVASSICVAT--- 754
 QY 1015 EPRTGVRGTGVLVEFSELAFLHRSFGDLTAEEKD-----QICDFLYGRVQAR 1062
 Db 755 -----TDVLAIEIQAL-----KGEVTVELKDSAFCTYILKSPKXICSLSHLTKWL 799
 QY 1063 ERQALARLR--TFQAFHVSVPSCPCLEQQDEERSTLKKLLGRYPPEERESQSPGGMED 1121
 Db 800 TEIESCKVRKLDIMSSAAVAISVNTSELSSGAKQTR--SLQGRIFDYFG-----D 850
 QY 1122 AQGPPEGARLQDWEDQ-----VRCDIROFLSLRPEEFSSRAVARIFHGIGSPCVPAQVY 1177
 Db 851 EKCSPPSKA-----TONCAFIRADIKVFLQSNROAKFTPRAIARIMGVSPAPFNSVW 904
 QY 1178 GQDRRFWRKYLHLSPHALVGLATEEL 1203
 Db 905 SK-THFWGRYMNVDPRVIMEAQTEL 929

RESULT 12
 Q7FAHO ORYSA
 ID Q7FAHO ORYSA PRELIMINARY; PRT; 874 AA.
 AC Q7FAHO7
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE QJ000223 09.17 protein.
 GN Name=QJ000223 09.17;
 OS Oryza sativa [japonica cultivar-group].
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriarthoideae; Oryzeae; Oryza.
 OC NCBI TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=23337377; PubMed=12447439; DOI=10.1038/nature01183;
 RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
 RA Weng Q., Zhang L., Lu Y., Wu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
 RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
 RA Lan L., Lai Y., Cheng Z., Gu M., Jang J., Li J., Hong G., Xue Y.,
 RA Han B.;
 RT "Sequence and analysis of rice chromosome 4.";
 RL Nature 420:316-320(2002).
 CC -!- FUNCTION: E1F4A is both a subunit of a high molecular weight
 CC protein complex involved in cap recognition and is required as a
 CC single polypeptide chain for mRNA binding to ribosome. It is an
 CC ATP-dependent single stranded DNA-binding protein with a sequence-
 CC independent unwinding activity (helicase) (By similarity).
 CC -!- SUBUNIT: e1f4p is a multi-subunit complex, the composition of
 CC which varies with external and internal environmental conditions.
 CC It is composed of at least E1F4A, E1F4E and E1F4G (By similarity).
 CC EMBL; AL606998; CAD41405.2; -; Genomic_DNA.
 DR Graneine; Q7FAHO; -;
 DR GO; GO:0005524; P-ATP binding; IEA.
 DR GO; GO:0008026; F-ATP-dependent helicase activity; IEA.
 DR GO; GO:0003677; F-DNA binding; IEA.
 DR GO; GO:0003743; F-translation initiation factor activity; IEA.
 DR GO; GO:0006259; P-DNA metabolism; IEA.
 DR GO; GO:0006412; P-protein biosynthesis; IEA.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR004589; RecQ.

DR Pfam; PF00270; DEAD; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR TIGRFAMs; TIGR00614; recQ_fam; 1.
 KW Helicase; Initiation factor; Protein biosynthesis.
 SQ SEQUENCE 874 AA; 96275 MW; 51B64C486617774F CRC64;
 Query Match 11.8%; Score 761; DB 2; Length 874;
 Best Local Similarity 23.8%; Pred. No. 2.9e-29;
 Matches 271; Conservative 163; Mismatches 377; Indels 330; Gaps 38;
 QY 110 KANLKGTLOAGPALGRPPWPLGRASSKASTPKP---GTGPVPSFAEKVSDPEPQLEPEQ 166
 Db 13 EGSLLSDVSASPPRRSP---PRPAPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 69
 QY 167 PRPRQLHLOASLSQRLSLDP--CWLORCHSEVDFLGA PKACRPDLGSEES---QLLIP 222
 Db 70 PAPPPTTLRAAALS-----DPHGLAAR-----IAAGSALTAAGTASSSSFRRLVQ 116
 QY 223 GESAVLFGAGSQGPASAFQEVSIKVGSPSSSGCKRRWNEEPWESPAQVOQESSQA 282
 Db 117 SRNPSFDATAFTAPASAPSEVPSAAPPPTTAA-----TDAPPQTPKRVHP 165
 QY 283 GPPSEGAGAVAEEDPPGEPVQAQPPQPCSPSPNRYHGLSPSSQARAGKAETAPLHIF 342
 Db 166 NSVSEVAAAAGAAEQP-----KRRGG----- 187
 QY 343 PRLARHGRNYVRLMKQKHYVRGRALRSLLRKQAWKQKWKKEGCEGCGGATVTTKES 402
 Db 188 -----SEGDFVRLNI--NGYGRRTTFKNSQAKSTKCRSWRRQ-----RAAGAT- 229
 QY 403 CFLNEQFDHMAAQCPRPASBEDTAVGPELVPSPQVPVPEVPSLDPTPLVLSLPGSQ 462
 Db 230 -----PRSGDEGLVAELERKQAASD-----SVLEAVE-----SV 264
 QY 463 AETPAEVQALEQL-----GHQAFRPGQERAVMRILSGISTLLVLTGAGKSLCYQLPAL 517
 Db 265 REDPSE--QNLKSLNAAHYGHSFRQGLEAIQIIVAGESTMLVLTGAGKSLCYQVDM 322
 QY 518 LYSRRSPCLTLVSPFLSLMDQVSGPLPPCLKAACIIHSGMTRKORESVLQIRAAQVHL 577
 Db 323 IL----PGLTLVSPFLSLMVDLRLKLPFAFPGGLLASSQTSDEFHDTLQRLRAGEIKVL 378
 QY 578 MLTPEALVGAGLPPAAQLPPVAFACIDEAHLCSQWSHNPFCVLRV--CKVLRERMGVHC 636
 Db 379 FVSPERFNEEFLLIFRDTLPILVALDEACISGWSHNPFCVLRVLRASLRKLNQVC 438
 QY 637 FLGTTATATRTASDVQAHLAVABEPDLHGPAVPPTNLHLSVSMRDTDOALLTLQGR 696
 Db 439 ILAMTATATTCTLEIMNNALEIPSD-NLIQTSQIRENLIQLSISTSDNRLKDLMLLKSP 497
 QY 697 FQNLDSIIYCNRRREDTERIAALLRTCLHAAWVPGSGRPAKTTAAAYHAGMCSRRRV 756
 Db 498 FVDMRSIIYCYCKFOAETDFVSKYL--CDN-----NITAKSYHSGLLIKNRGRV 543
 QY 757 QRAFWQGLRVVATVAFMGMLDRPDVRAVLHLGLPPSFESVQAVGRAGRDGPAHCHL 816
 Db 544 QELFCSNKRIVF-----SF-----NF 559
 QY 817 FLQPGEDLRLRRHVHADSTDFLAKRLVORVFA-----CTCTCTPPSPSEGAAGVE 871
 Db 560 TEQP-----FLGSDGVGYAMSKFLQIFSENTTGCICSLAK----- 597
 QY 872 RPVPKYPQEAQLSHQAAPGRRVCMGHERALPIQLTVQALDMPERAIETLLCYLELHP 931
 Db 598 -----ELTSRKFDIKEEVLLTILTLQLEIGD 622
 QY 932 HHWLELLATTYTHCELNCPGPAQLOALAHRCPPPLAVC-----LAQQLPDPQSSSVF 987
 Db 623 QQYIKLLPQFSVTCTL-----YFHKTSIPOLLADKDLILRSVLNRSEMKDGHVYVF 671
 QY 988 DMVKLVDSMGWELASVRRALCOLQWHDHPRGTGVRRTGVLGVVEFS--ELAFHLRSPG----- 1041
 Db 672 DIPRIANDLKITMNEV-----FDHLHK-----LKFSGEISFELKDPAYCYVI 713

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA], AND SEQUENCE REVISION
 TO 254.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RN Science 277:1453-1474 (1997).
 [4]
 RP SEQUENCE REVISION TO 502-503.
 RA Horluch T., Kousuge T., Mori H., Perna N.T., Plunkett G. III,
 RA Riley M., Rudd K.E., Serres M.H., Thomas G.H., Wanner B.L.;
 RT "Workshop on annotation of *Escherichia coli* K-12.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 [5]
 RP FUNCTION, AND PROTEIN SEQUENCE OF 1-5.
 RX MEDLINE=90311913; PubMed=2164680;
 RA Umez K., Nakayama K., Nakayama H.;
 RA "Escherichia coli RecQ protein is a DNA helicase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5363-5367 (1990).
 CC -!- FUNCTION: Involved in the recF recombination pathway; its gene
 CC expression is under the regulation of the SOS system. It is a DNA
 CC helicase.
 CC -!- INTERACTION:
 CC P06958:aceF; NbExp=1; IntAct=EBI-552058, EBI-542683;
 CC P06959:aceF; NbExp=1; IntAct=EBI-552058, EBI-542707;
 CC P28905:holC; NbExp=1; IntAct=EBI-552058, EBI-549159;
 CC P28632:holD; NbExp=1; IntAct=EBI-552058, EBI-549176;
 CC P00391:lpdA; NbExp=1; IntAct=EBI-552058, EBI-542856;
 CC P0A7Y4:rmhA; NbExp=1; IntAct=EBI-552058, EBI-556910;
 CC P02370:tpnB; NbExp=1; IntAct=EBI-552058, EBI-543024;
 CC P02370:tpnB; NbExp=1; IntAct=EBI-552058, EBI-543971;
 CC P28043:rsbB; NbExp=1; IntAct=EBI-552058, EBI-549252;
 CC P06612:topA; NbExp=1; IntAct=EBI-552058, EBI-544172;
 CC P14294:topB; NbExp=1; IntAct=EBI-552058, EBI-552080;
 CC -!- SIMILARITY: Belongs to the helicase family. RecQ subfamily.
 CC -!- SIMILARITY: Contains 1 HRDC domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; M30198; AAA24517.1; ALT_INIT; Genomic_DNA.
 CC EMBL; M87049; AAA67618.1; ALT_INIT; Genomic_DNA.
 CC EMBL; U00096; AAT48221.1; -; Genomic_DNA.
 CC PDB; 1OYW; X-ray; A=1-522.
 CC PDB; 1OYV; X-ray; A=1-522.
 CC IntAct; P15043; -;
 CC EcoBASE; EB0826; -;
 CC EcoGene; EG10833; recQ.
 CC InterPro; IPR006293; ATP RecQ.
 CC InterPro; IPR001410; DEAD.
 CC InterPro; IPR011545; DEAD/DEAH_N.
 CC InterPro; IPR002464; DEAH/DEAH_N.
 CC InterPro; IPR001650; Helicase_C.
 CC InterPro; IPR002121; HRDC.
 CC InterPro; IPR004589; RecQ.
 CC PANTHER; PTHR13710; RecQ; 1.
 CC Pfam; PF00270; DEAD; 1.
 CC Pfam; PF00271; Helicase_C; 1.
 CC Pfam; PF00570; HRDC; 1.
 CC SMART; SM00487; DEXDC; 1.
 CC SMART; SM00490; HELIC; 1.
 CC SMART; SM00341; HRDC; 1.
 CC TIGRFAMs; TIGR01389; recQ; 1.
 CC TIGRFAMs; TIGR00614; recQ_fam; 1.
 CC PROSITE; PS00690; DEAH_ATP_HELICASE; FALSE_NEG.
 CC PROSITE; PSS0967; HRDC; 1.

KW 3D-structure; ATP-binding; Complete proteome;
 KW Direct protein sequencing; DNA damage; DNA recombination; DNA repair;
 KW DNA-binding; Helicase; Hydrolase; Nucleotide-binding; SOS response.
 FT INIT MET 0
 FT DOMAIN 528 608 HRDC.
 FT NP BIND 46 53 ATP (Potential).
 FT MOTIF 145 148 DEAH box.
 FT CONFLICT 254 254 R -> A (in Ref. 1 and 2).
 FT CONFLICT 502 503 RG -> A (in Ref. 1 and 2).
 FT CONFLICT 599 599 R -> P (in Ref. 1).
 SQ SEQUENCE 608 AA; 68232 MW; AA01654270061AE1 CRC64;
 Query Match 9.8%; Score 631.5; DB 1; Length 608;
 Best Local Similarity 37.6%; Pred. No. 4.4e-23;
 Matches 164; Conservative 63; Mismatches 134; Indels 75; Gaps 15;
 Qy 474 EQLGHQAFRPGQRAVRNRIISGLISTLLVPTGAGKSLCYQLPALLYRRSPCLFLVVSPL 533
 Db 18 ETFGYQQFRPQEEIITVLSGRDCLVVMPTGGKSLCYQIPALLNG---LTVVVSPL 73
 Qy 534 LSLMDDQVGLPP-CLKACIHSQMTKQRESVLQKTRAAQVHVMLTPEALVGAGGLPP 592
 Db 74 ISLMKDQVQLQANGVAAACLNSTQTRQQLVMTGCTGQIRLLYIAPERMLDNLFEH 133
 Qy 593 AAQLPPVAFACIDEAGCLSQSHNFRPCYLRVCKVLRERMGVHCFGLGTATATRTASDV 652
 Db 134 LAHNPVLLA-VDEAHCIQGHDFRPEYAAALGQ-LKQRPFTLPFMALTATADTTTQDI 191
 Qy 653 AQHLAVAEPLDHGPAFVPTNHLISVSNDRDITDQALLTLQGRFONLDSI----- 703
 Db 192 VRLLG-----LNDPL-----IQIS-SFDRPNIRVLMW---EKFKPLDQLMRVVOEQRG 235
 Qy 704 ---IYCNRRDTERIAALLTCLHAAVPGSGRPAKPTTAEYAHGMCRRERRVQRAF 760
 Db 236 KSGIYCNRAKVEDTAARL-----QSKGISAAAYHAGLENNVRADVQSKF 281
 Qy 761 MQQLRVVAVTAVRGMLDRPDVRAVLHGLGPPSFESVYQAVGRAGRDGQPAHCHLFL- 818
 Db 282 QRDDLQIVVATVAFGMGINKPNVFPVHFDPRIPTIESYQETGRAGRDGLPFAEMLFYDP 341
 Qy 819 -----QPGEDLRELRRH---VHADSTDFLAVKRLV-----QRVFPACTC- 855
 Db 342 ADMAWLRCLKEKQGG-LQDIERHKLNAMGAFAEAQTCRLVLLNYFEGRQEGPCGNG 400
 Qy 856 TCTRPPSPQEGAVGGE 871
 Db 401 ICLDPPKQYDGGSTDAQ 416
 RESULT 15
 Q8X8N1_ECO57
 ID Q8X8N1_ECO57 PRELIMINARY; PRT; 611 AA.
 AC Q8X8N1_07A9D7;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
 DE ATP-dependent DNA helicase.
 GN Name=recQ; OrderedLocusNames=EC64752, z5343;
 OS *Escherichia coli* O157:H7;
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OX NCBI_TaxID=83334;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:43:16 ; Search time 52 Seconds
(without alignments)
1920.620 Million cell updates/sec

Title: US-09-889-325-4

Perfect score: 6424
Sequence: 1 MERLDVRLQAWERAFRR.....HLSFHALVGLATELLOVAR 1208

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCRTUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 6424 | 100.0 | 1208 | 2 | US-09-463-702A-2 |
| 2 | 6424 | 100.0 | 1208 | 2 | US-09-699-135-2 |
| 3 | 1883 | 29.3 | 361 | 2 | US-09-463-702A-37 |
| 4 | 1883 | 29.3 | 361 | 2 | US-09-699-135-37 |
| 5 | 632.5 | 9.8 | 607 | 2 | US-08-781-891-75 |
| 6 | 632.5 | 9.8 | 607 | 2 | US-09-618-166-75 |
| 7 | 621 | 9.7 | 370 | 1 | US-08-559-303B-77 |
| 8 | 621 | 9.7 | 370 | 2 | US-08-175-828-77 |
| 9 | 621 | 9.7 | 370 | 2 | US-09-753-143-77 |
| 10 | 613.5 | 9.6 | 334 | 2 | US-09-463-702A-38 |
| 11 | 613.5 | 9.6 | 334 | 2 | US-09-699-135-38 |
| 12 | 606 | 9.4 | 619 | 2 | US-09-489-039A-8075 |
| 13 | 591.5 | 9.2 | 641 | 2 | US-09-543-681A-4359 |
| 14 | 585 | 9.1 | 1394 | 2 | US-09-902-540-16497 |
| 15 | 554.5 | 8.6 | 589 | 2 | US-09-134-000C-5710 |
| 16 | 547 | 8.5 | 849 | 2 | US-09-252-991A-17953 |
| 17 | 545 | 8.5 | 106 | 2 | US-09-463-702A-43 |
| 18 | 545 | 8.5 | 106 | 2 | US-09-699-135-43 |
| 19 | 529 | 8.2 | 616 | 2 | US-09-107-532A-7258 |
| 20 | 526 | 8.2 | 637 | 2 | US-08-134-001C-5658 |
| 21 | 525.5 | 8.2 | 1328 | 2 | US-08-781-891-76 |
| 22 | 525.5 | 8.2 | 1328 | 2 | US-09-618-166-76 |
| 23 | 521 | 8.1 | 473 | 2 | US-09-107-532A-5941 |
| 24 | 506.5 | 7.9 | 410 | 2 | US-09-949-016-9327 |
| 25 | 494 | 7.7 | 355 | 2 | US-09-902-540-15688 |
| 26 | 487 | 7.6 | 393 | 1 | US-08-559-303B-76 |
| 27 | 487 | 7.6 | 393 | 2 | US-09-175-828-76 |

| | | | | | | |
|----|-------|-----|------|---|----------------------|-------------------|
| 28 | 487 | 7.6 | 393 | 2 | US-09-753-143-76 | Sequence 76, Appl |
| 29 | 485.5 | 7.6 | 1401 | 2 | US-09-127-670-6 | Sequence 6, Appli |
| 30 | 481.5 | 7.5 | 333 | 2 | US-09-463-702A-39 | Sequence 39, Appl |
| 31 | 481.5 | 7.5 | 333 | 2 | US-09-699-135-39 | Sequence 39, Appl |
| 32 | 479.5 | 7.5 | 1401 | 2 | US-08-781-891-206 | Sequence 206, App |
| 33 | 479.5 | 7.5 | 1401 | 2 | US-09-618-166-206 | Sequence 206, App |
| 34 | 478.5 | 7.4 | 330 | 2 | US-09-134-000C-4772 | Sequence 4772, Ap |
| 35 | 476 | 7.4 | 1409 | 2 | US-09-949-001-22 | Sequence 22, Appl |
| 36 | 475 | 7.4 | 1269 | 2 | US-08-781-891-74 | Sequence 74, Appl |
| 37 | 475 | 7.4 | 1269 | 2 | US-09-618-166-74 | Sequence 71, Appl |
| 38 | 475 | 7.4 | 1432 | 2 | US-08-781-891-71 | Sequence 71, Appl |
| 39 | 475 | 7.4 | 1432 | 2 | US-09-618-166-71 | Sequence 16, Appl |
| 40 | 475 | 7.4 | 1432 | 2 | US-09-949-001-16 | Sequence 20040, A |
| 41 | 473 | 7.4 | 879 | 2 | US-09-248-796A-20040 | Sequence 75, Appl |
| 42 | 464.5 | 7.2 | 386 | 1 | US-08-559-303B-75 | Sequence 75, Appl |
| 43 | 464.5 | 7.2 | 386 | 2 | US-09-175-828-75 | Sequence 75, Appl |
| 44 | 464.5 | 7.2 | 386 | 2 | US-09-753-143-75 | Sequence 75, Appl |
| 45 | 464.5 | 7.2 | 659 | 2 | US-08-781-891-77 | Sequence 77, Appl |

ALIGNMENTS

RESULT 1
US-09-463-702A-2
; Sequence 2, Application US/09463702A
; Patent No. 6335435
; GENERAL INFORMATION:
; APPLICANT: AGENE Research Institute, Co., Ltd.
; APPLICANT: HIRAKI AND ASSOCIATES
; APPLICANT: SHIMAMOTO, AKIRO
; APPLICANT: KITAO, SAORI
; APPLICANT: FURUICHI, YASUHIRO
; FILE OF INVENTION: HIRAI150
; TITLE OF INVENTION: HUMAN GENE RECO4 ENCODING HELICASE
; CURRENT APPLICATION NUMBER: US/09/463,702A
; CURRENT FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/JP98/03114
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: JAPAN 9/200387
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-463-702A-2

| | | | | |
|-----------------------|--------------|--|------------|--------------|
| Query Match | 100.0% | Score 6424; | DB 2; | Length 1208; |
| Best Local Similarity | 100.0% | Pred. No. 0; | | |
| Matches 1208; | Conservative | 0; | Mismatches | 0; |
| Indels | 0; | Gaps | 0; | |
| Qy | 1 | MERLDVRLQAWERAFRRQRORRRPSQDDVEAAPEETRALYREYRTLKRTTGOAGGGLR | 60 | |
| Db | 1 | MERLDVRLQAWERAFRRQRORRRPSQDDVEAAPEETRALYREYRTLKRTTGOAGGGLR | 60 | |
| Qy | 61 | SSESLPAAAEBAEPRCWCWPHLNRAATKSPQTPGCRSQSGVDPDYQBLKANLKTLOAG | 120 | |
| Db | 61 | SSESLPAAAEBAEPRCWCWPHLNRAATKSPQTPGCRSQSGVDPDYQBLKANLKTLOAG | 120 | |
| Qy | 121 | PALGRRPWLGRASSKASTPKPGTGPVPSFAEKVSDPEPQLPEPQPPRGLHLOASLS | 180 | |
| Db | 121 | PALGRRPWLGRASSKASTPKPGTGPVPSFAEKVSDPEPQLPEPQPPRGLHLOASLS | 180 | |
| Qy | 181 | ORLGSLDPGLQRCHSEVPDFLGAPKACRPDLGSESOLLIPGESAVLPGAGSGQGPAS | 240 | |
| Db | 181 | ORLGSLDPGLQRCHSEVPDFLGAPKACRPDLGSESOLLIPGESAVLPGAGSGQGPAS | 240 | |
| Qy | 241 | AFQEVSIKVGSPQPSSSGGKKRRWNEEPWESPAQVQQESSQAGPPSEGAAGVAVEEDPPG | 300 | |
| Db | 241 | AFQEVSIKVGSPQPSSSGGKKRRWNEEPWESPAQVQQESSQAGPPSEGAAGVAVEEDPPG | 300 | |

```
Qy 301 EPVQAQPPQPCSSPNRYHGLSPSSQARAKAGTAPLHPPLARHDRGNVYRLNMKQ 360
Db 301 EPVQAQPPQPCSSPNRYHGLSPSSQARAKAGTAPLHPPLARHDRGNVYRLNMKQ 360
Qy 361 KHYVGRALRSRLRKQAWKQWKKGECFGGGATVTTKESCFLENEQFDHWAACQPRPA 420
Db 361 KHYVGRALRSRLRKQAWKQWKKGECFGGGATVTTKESCFLENEQFDHWAACQPRPA 420
Qy 421 SEEDTDAVGPEPLVPSPQVPEVPSLQPTVPLYSLSGSPGQLAETPAEVFOALQHQHQA 480
Db 421 SEEDTDAVGPEPLVPSPQVPEVPSLQPTVPLYSLSGSPGQLAETPAEVFOALQHQHQA 480
Qy 481 FRPQGERAVMRILSGISTLLVLPFGAGKSLCYQLPALLYSRSPCLTLVWSPILLSMDQ 540
Db 481 FRPQGERAVMRILSGISTLLVLPFGAGKSLCYQLPALLYSRSPCLTLVWSPILLSMDQ 540
Qy 541 VSGLPCLKAACIHSQWTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPPAAQLPPVA 600
Db 541 VSGLPCLKAACIHSQWTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPPAAQLPPVA 600
Qy 601 FACIDEAHCLSQWSHNFRPCYLRCVKLRERMVGHCFGLGTATATRTTASDVAAQLHVAE 660
Db 601 FACIDEAHCLSQWSHNFRPCYLRCVKLRERMVGHCFGLGTATATRTTASDVAAQLHVAE 660
Qy 661 EPDLHGPAVPVPTNLHLVSMDRDTQALLTLQKRFQNLDSIIYCNREDTERIAALL 720
Db 661 EPDLHGPAVPVPTNLHLVSMDRDTQALLTLQKRFQNLDSIIYCNREDTERIAALL 720
Qy 721 RTCLHAHVPGSGGRAPKTATAEAHAGMCSRERRVQAFMOGQLRVVVAATVAFGMGLDR 780
Db 721 RTCLHAHVPGSGGRAPKTATAEAHAGMCSRERRVQAFMOGQLRVVVAATVAFGMGLDR 780
Qy 781 PDVRAVLHLGLPPSPFESVQAVGRAGRDGPAAHCHLFLQPGEDLRELRRHVHADSTDFL 840
Db 781 PDVRAVLHLGLPPSPFESVQAVGRAGRDGPAAHCHLFLQPGEDLRELRRHVHADSTDFL 840
Qy 841 AVKGLVQRVFPACTCTCTRPSPSEGA VGGERPVPKYPPOBAEQLSHQAAAPGPRRVCMGH 900
Db 841 AVKGLVQRVFPACTCTCTRPSPSEGA VGGERPVPKYPPOBAEQLSHQAAAPGPRRVCMGH 900
Qy 901 ERALPIQLTVOALDMPBEATETLLCYLELHPHWHLELLATYTHCLNCPGPAQLQALA 960
Db 901 ERALPIQLTVOALDMPBEATETLLCYLELHPHWHLELLATYTHCLNCPGPAQLQALA 960
Qy 961 HRCPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPTGV 1020
Db 961 HRCPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPTGV 1020
Qy 1021 RRGTVLVEFSELAFHLRSPGDLTAEBKQIQICDFLYGRVQARERQALRLRRTFOAFHSV 1080
Db 1021 RRGTVLVEFSELAFHLRSPGDLTAEBKQIQICDFLYGRVQARERQALRLRRTFOAFHSV 1080
Qy 1081 AFPCSGCLQDEERSTRKDLGRYFEEBEGQEPGEMDAQEPGQARLQWEDQVR 1140
Db 1081 AFPCSGCLQDEERSTRKDLGRYFEEBEGQEPGEMDAQEPGQARLQWEDQVR 1140
Qy 1141 CDIRQFLSLRPEEFSSRAVARIFPHIGSPCYPAQVYQDQRRFRWKYHLHSFHALVGLAT 1200
Db 1141 CDIRQFLSLRPEEFSSRAVARIFPHIGSPCYPAQVYQDQRRFRWKYHLHSFHALVGLAT 1200
Qy 1201 BELLQVAR 1208
Db 1201 BELLQVAR 1208
```

RESULT 2

US-09-699-135-2

; Sequence 2, Application US/09699135

; Patent No. 6472513

; GENERAL INFORMATION:

; APPLICANT: AGENCE Research Institute, Co., Ltd.

; APPLICANT: HIRAKI AND ASSOCIATES

; APPLICANT: SHIMAMOTO, AKIRO

```
; APPLICANT: KITAO, SAORI
; APPLICANT: FURUICHI, YASUHIRO
; TITLE OF INVENTION: HUMAN GENE REQ4 ENCODING HELICASE
; FILE REFERENCE: HIRAI150
; CURRENT APPLICATION NUMBER: US/09/699,135
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US/09/463,702A
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/JP98/03114
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: JAPAN 9/200387
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-699-135-2
```

Query Match 100.0%; Score 6424; DB 2; Length 1208;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MERLRDVRERLQAWERAFRRQGRRRPSQDDVEAAPETRALYREYRTLKRTTGQAGGLR 60
Db 1 MERLRDVRERLQAWERAFRRQGRRRPSQDDVEAAPETRALYREYRTLKRTTGQAGGLR 60
Qy 61 SSESLLPAAAEAEPEPCWGHILNRAATKSPQTPGRSRQGSVDYQGRLLKANLKGTLQAG 120
Db 61 SSESLLPAAAEAEPEPCWGHILNRAATKSPQTPGRSRQGSVDYQGRLLKANLKGTLQAG 120
Qy 121 PALGRRPWPLGRASSKASTPKPGTGPVPSPFAKVSDEPPQLPEPOPGRLOHLOASLS 180
Db 121 PALGRRPWPLGRASSKASTPKPGTGPVPSPFAKVSDEPPQLPEPOPGRLOHLOASLS 180
Qy 181 QRLGSLDPCWLQCHSEVPDFLGAPKACRDPDLGSEESQLLIPGESAVLPGAGSQGPEAS 240
Db 181 QRLGSLDPCWLQCHSEVPDFLGAPKACRDPDLGSEESQLLIPGESAVLPGAGSQGPEAS 240
Qy 241 AFEVSIIRVGSPOPPSSSGGKRRNNEPWPSPQAQVQOESSQAQPPSPGAGAVAEEDPPG 300
Db 241 AFEVSIIRVGSPOPPSSSGGKRRNNEPWPSPQAQVQOESSQAQPPSPGAGAVAEEDPPG 300
Qy 301 EPVQAQPPQPCSSPNRYHGLSPSSQARAKAGTAPLHPPLARHDRGNVYRLNMKQ 360
Db 301 EPVQAQPPQPCSSPNRYHGLSPSSQARAKAGTAPLHPPLARHDRGNVYRLNMKQ 360
Qy 361 KHYVGRALRSRLRKQAWKQWKKGECFGGGATVTTKESCFLENEQFDHWAACQPRPA 420
Db 361 KHYVGRALRSRLRKQAWKQWKKGECFGGGATVTTKESCFLENEQFDHWAACQPRPA 420
Qy 421 SEEDTDAVGPEPLVPSPQVPEVPSLQPTVPLYSLSGSPGQLAETPAEVFOALQHQHQA 480
Db 421 SEEDTDAVGPEPLVPSPQVPEVPSLQPTVPLYSLSGSPGQLAETPAEVFOALQHQHQA 480
Qy 481 FRPQGERAVMRILSGISTLLVLPFGAGKSLCYQLPALLYSRSPCLTLVWSPILLSMDQ 540
Db 481 FRPQGERAVMRILSGISTLLVLPFGAGKSLCYQLPALLYSRSPCLTLVWSPILLSMDQ 540
Qy 541 VSGLPCLKAACIHSQWTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPPAAQLPPVA 600
Db 541 VSGLPCLKAACIHSQWTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPPAAQLPPVA 600
Qy 601 FACIDEAHCLSQWSHNFRPCYLRCVKLRERMVGHCFGLGTATATRTTASDVAAQLHVAE 660
Db 601 FACIDEAHCLSQWSHNFRPCYLRCVKLRERMVGHCFGLGTATATRTTASDVAAQLHVAE 660
Qy 661 EPDLHGPAVPVPTNLHLVSMDRDTQALLTLQKRFQNLDSIIYCNREDTERIAALL 720
Db 661 EPDLHGPAVPVPTNLHLVSMDRDTQALLTLQKRFQNLDSIIYCNREDTERIAALL 720
Qy 721 RTCLHAHVPGSGGRAPKTATAEAHAGMCSRERRVQAFMOGQLRVVVAATVAFGMGLDR 780
```

Db 721 RTCLHAAVPPGSGGRAPKTTAEAYHAGMCSRRRRVQRAFMQGLRVVVVATVAFGMGLDR 780
 Qy 781 PDVRAVLHGLPPSPFESVQAVGRAGRDGQPAHCHLFLQPGEDLRELRRHVHADSTDFL 840
 Db 781 PDVRAVLHGLPPSPFESVQAVGRAGRDGQPAHCHLFLQPGEDLRELRRHVHADSTDFL 840
 Qy 841 AVKGLVQRVPACTCTCTRPSPSEBEGAVGGRPVKYPQPAEQLSHQAAPGPRVCMGH 900
 Db 841 AVKGLVQRVPACTCTCTRPSPSEBEGAVGGRPVKYPQPAEQLSHQAAPGPRVCMGH 900
 Qy 901 BRALPIQTLVQALDMPERAEITLLCYLELHPHMLELLATTYTHCRNLNCPGPGPAQLQALA 960
 Db 901 BRALPIQTLVQALDMPERAEITLLCYLELHPHMLELLATTYTHCRNLNCPGPGPAQLQALA 960
 Qy 961 HRCPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDEHPRTGV 1020
 Db 961 HRCPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDEHPRTGV 1020
 Qy 1021 RRGTVLVEFSELAPHLRSPGDLTAEEKDQICDFLYGRVQARERQALRLRTTQAFHSV 1080
 Db 1021 RRGTVLVEFSELAPHLRSPGDLTAEEKDQICDFLYGRVQARERQALRLRTTQAFHSV 1080
 Qy 1081 APPSCGCLQOODEERSFELKDLGRYFEEREGQEPGGMEDAQGPQOARLQWEDQVR 1140
 Db 1081 APPSCGCLQOODEERSFELKDLGRYFEEREGQEPGGMEDAQGPQOARLQWEDQVR 1140
 Qy 1141 CDIROFLSLRPEEKFSRAVARIFHGIGSPCYPAQVYGQDRFRFWRKYLHLSFHALVGLAT 1200
 Db 1141 CDIROFLSLRPEEKFSRAVARIFHGIGSPCYPAQVYGQDRFRFWRKYLHLSFHALVGLAT 1200
 Qy 1201 BELLQVAR 1208
 Db 1201 BELLQVAR 1208

RESULT 3

US-09-463-702A-37
 ; Sequence 37, Application US/09463702A
 ; Patent No. 6335435
 ; GENERAL INFORMATION:
 ; APPLICANT: AGENE Research Institute, Co., Ltd.
 ; APPLICANT: HIRAKI AND ASSOCIATES
 ; APPLICANT: SHIMAMOTO, AKIRO
 ; APPLICANT: KITAO, SAORI
 ; APPLICANT: FURUICHI, YASUHIRO
 ; TITLE OF INVENTION: HUMAN GENE RECO4 ENCODING HELICASE
 ; FILE REFERENCE: HIRAI150
 ; CURRENT APPLICATION NUMBER: US/09/463, 702A
 ; PRIOR FILING DATE: 2000-01-24
 ; PRIOR APPLICATION NUMBER: PCT/JP98/03114
 ; PRIOR FILING DATE: 1998-07-10
 ; PRIOR APPLICATION NUMBER: JAPAN 9/200387
 ; PRIOR FILING DATE: 1997-07-25
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 37
 ; LENGTH: 361
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-463-702A-37

Query Match 29.3%; Score 1883; DB 2; Length 361;
 Best Local Similarity 100.0%; Pred. No. 2.4e-147;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 QALEQLGHOAFRPGQERAVMRILSGISTLLVPTGAGKSLCYQLPALLYSRSPCLTLVV 530
 Db 1 QALEQLGHOAFRPGQERAVMRILSGISTLLVPTGAGKSLCYQLPALLYSRSPCLTLVV 60
 Qy 531 SPLLSLMDQVSGLPCLKAACHSGMTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGL 590
 Db 61 SPLLSLMDQVSGLPCLKAACHSGMTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGL 120

Qy 591 PPAQLPPVAFACIDEAHCLSQSHNFRPCYLVRCKVLRRMGVHCFGLGTATATRTAS 650
 Db 121 PPAQLPPVAFACIDEAHCLSQSHNFRPCYLVRCKVLRRMGVHCFGLGTATATRTAS 180
 Qy 651 DVAQHLLAVAEEDPLHGPAPVPTNLHLSVSMRDRTDQALLTLQGRFQNLDSIIYCNRR 710
 Db 181 DVAQHLLAVAEEDPLHGPAPVPTNLHLSVSMRDRTDQALLTLQGRFQNLDSIIYCNRR 240
 Qy 711 EDTERIAALLRCTLHAAWVPGSGGRAPKTTAEAYHAGMCSRRRRVQRAFMQGLRVVVA 770
 Db 241 EDTERIAALLRCTLHAAWVPGSGGRAPKTTAEAYHAGMCSRRRRVQRAFMQGLRVVVA 300
 Qy 771 TVAFGMGLDRPDVRAVLHGLPPSFESYVOAVGRAGRDGQPAHCHLFLQPGEDLRELRR 830
 Db 301 TVAFGMGLDRPDVRAVLHGLPPSFESYVOAVGRAGRDGQPAHCHLFLQPGEDLRELRR 360
 Qy 831 H 831
 Db 361 H 361

RESULT 4

US-09-699-135-37
 ; Sequence 37, Application US/09699135
 ; Patent No. 6472513
 ; GENERAL INFORMATION:
 ; APPLICANT: AGENE Research Institute, Co., Ltd.
 ; APPLICANT: HIRAKI AND ASSOCIATES
 ; APPLICANT: SHIMAMOTO, AKIRO
 ; APPLICANT: KITAO, SAORI
 ; APPLICANT: FURUICHI, YASUHIRO
 ; TITLE OF INVENTION: HUMAN GENE RECO4 ENCODING HELICASE
 ; FILE REFERENCE: HIRAI150
 ; CURRENT APPLICATION NUMBER: US/09/699,135
 ; PRIOR FILING DATE: 2000-10-27
 ; PRIOR APPLICATION NUMBER: US/09/463, 702A
 ; PRIOR FILING DATE: 2000-01-24
 ; PRIOR APPLICATION NUMBER: PCT/JP98/03114
 ; PRIOR FILING DATE: 1998-07-10
 ; PRIOR APPLICATION NUMBER: JAPAN 9/200387
 ; PRIOR FILING DATE: 1997-07-25
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 37
 ; LENGTH: 361
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-699-135-37

Query Match 29.3%; Score 1883; DB 2; Length 361;
 Best Local Similarity 100.0%; Pred. No. 2.4e-147;
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 QALEQLGHOAFRPGQERAVMRILSGISTLLVPTGAGKSLCYQLPALLYSRSPCLTLVV 530
 Db 1 QALEQLGHOAFRPGQERAVMRILSGISTLLVPTGAGKSLCYQLPALLYSRSPCLTLVV 60
 Qy 531 SPLLSLMDQVSGLPCLKAACHSGMTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGL 590
 Db 61 SPLLSLMDQVSGLPCLKAACHSGMTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGL 120
 Qy 591 PPAQLPPVAFACIDEAHCLSQSHNFRPCYLVRCKVLRRMGVHCFGLGTATATRTAS 650
 Db 121 PPAQLPPVAFACIDEAHCLSQSHNFRPCYLVRCKVLRRMGVHCFGLGTATATRTAS 180
 Qy 651 DVAQHLLAVAEEDPLHGPAPVPTNLHLSVSMRDRTDQALLTLQGRFQNLDSIIYCNRR 710
 Db 181 DVAQHLLAVAEEDPLHGPAPVPTNLHLSVSMRDRTDQALLTLQGRFQNLDSIIYCNRR 240
 Qy 711 EDTERIAALLRCTLHAAWVPGSGGRAPKTTAEAYHAGMCSRRRRVQRAFMQGLRVVVA 770
 Db 241 EDTERIAALLRCTLHAAWVPGSGGRAPKTTAEAYHAGMCSRRRRVQRAFMQGLRVVVA 300

Db 74 ISLMKQVQDQVQANGVAAACLNSTQREQQLEVMVTCGTGQIRLLYIAPERLMLDNFLEH 133
Qy 593 AAQLPPVAFACIDEAHCLSQSHNFRPCYLKRVCKVLRMRMGVHCFGLGTATATARTASDV 652
Db 134 LAHNVPVLLA-VDEAHCSIQMGHDFRPEYALGQ-LRQRPPTLPFMALTATADDTTRQDI 191
Qy 653 AQHLAVAEEDPLHGPAPVPTNLHLVSVMDRDTDQALLTLLOGKRFQNLDSI----- 703
Db 192 VRLIG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKPKLDQLMRYVQORQ 235
Qy 704 ---IICNRREDTERIAALLRTCLHAAWVPGSGGRAPKTTAEAYHAGMCSRERRVORAF 760
Db 236 KSGIYCNSRAKVEDTAAAL-----QSGISAAAYHAGLENNVRADVQKPF 281
Qy 761 MQQLRVVVVATVAFGMGLDRPDVRAVLHLGLPPSFESYVQAVGRAGRDGQPAHCHLFL-- 818
Db 282 QRDDLQIVVATVAFGMGINKENPVFVHFDPRIENIESYVQETGRAGRDGLPAEAMLFYDP 341
Qy 819 -----QPGEDLRELRRH---VHADSTDFLAVKELV-----QRVPFACCTC- 855
Db 342 ADMAWLRCLCEKPPQGO-LQDIERHKLNAMGAFEAOTCRRLVLLNLYFEGRQBPCCGND 400
Qy 856 TCTRPSPQEGAVGGE 871
Db 401 ICLDPPKQYDGSDDAQ 416

RESULT 7

US-08-559-303B-77
; Sequence 77, Application US/08559303B
; Patent No. 5824501
; GENERAL INFORMATION:
; APPLICANT: NATHAN A. ELLIS, JAMES GERMAN, AND JOANNA
; APPLICANT: GRODEN
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BLOOM'S SYNDROME
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
; STREET: 90 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,303B
; FILING DATE: NOVEMBER 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ELIZABETH A. BOGOSIAN
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/65
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: OTHER NUCLEIC ACID
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY:
; LOCATION:

; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; US-08-559-303B-77
Query Match 9.7%; Score 621; DB 1; Length 370;
Best Local Similarity 40.0%; Pred. No. 7.1e-43;
Matches 154; Conservative 57; Mismatches 110; Indels 64; Gaps 12;
Qy 474 EQLGHQAFRRGQERAVNRILSGISTLLVLTGTAGKSLCYQLPALLYRRRPFCLTLVVVSP 533
Db 4 ETEGYOQFRPQGEIITVLSGRDCLVVMPTGGKSLCYQIPALLNG-----LTVVVVSP 59
Qy 534 LSLMDDDOVGLPP-CLKAAACHSGMTRKQRESVLQKIRAAQVHVLMLTPEALVAGAGLPP 592
Db 60 ISLMKQVQDQVQANGVAAACLNSTQREQQLEVMVTCGTGQIRLLYIAPERLMLDNFLEH 119
Qy 593 AAQLPPVAFACIDEAHCLSQSHNFRPCYLKRVCKVLRMRMGVHCFGLGTATATARTASDV 652
Db 120 LAHNVPVLLA-VDEAHCSIQMGHDFRPEYALGQ-LRQRPPTLPFMALTATADDTTRQDI 177
Qy 653 AQHLAVAEEDPLHGPAPVPTNLHLVSVMDRDTDQALLTLLOGKRFQNLDSI----- 703
Db 178 VRLIG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKPKLDQLMRYVQORQ 221
Qy 704 ---IICNRREDTERIAALLRTCLHAAWVPGSGGRAPKTTAEAYHAGMCSRERRVORAF 760
Db 222 KSGIYCNSRAKVEDTAAAL-----QSGISAAAYHAGLENNVRADVQKPF 267
Qy 761 MQQLRVVVVATVAFGMGLDRPDVRAVLHLGLPPSFESYVQAVGRAGRDGQPAHCHLFL-- 818
Db 268 QRDDLQIVVATVAFGMGINKENPVFVHFDPRIENIESYVQETGRAGRDGLPAEAMLFYDP 327
Qy 819 -----QPGEDLRELRRH 831
Db 328 ADMAWLRCLCEKPPQGO-LQDIERH 351
RESULT 8
US-09-175-828-77
; Sequence 77, Application US/09175828
; Patent No. 6221643
; GENERAL INFORMATION:
; APPLICANT: NATHAN A. ELLIS, JAMES GERMAN, AND JOANNA
; APPLICANT: GRODEN
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BLOOM'S SYNDROME
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
; STREET: 90 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/175,828
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,303
; FILING DATE: NOVEMBER 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ELIZABETH A. BOGOSIAN
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/65
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766

```
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: OTHER NUCLEIC ACID
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
;
US-09-175-828-77

Query Match          9.7%; Score 621; DB 2; Length 370;
Best Local Similarity 40.0%; Pred. No. 7.1e-43;
Matches 154; Conservative 57; Mismatches 110; Indels 64; Gaps 12;

QY 474 EQLGHOAFRPGQRAVNRILSGISTLLVPTGAGKSLCYQLPALLYRRSPCLTLVVSPL 533
DB 4 ETFGYQFRPGQEEIITVLSGRDCLVVMPTGGKSLCYQIPALLNG----LTVVVSPL 59

QY 534 LSLMDDQVSGLPP-CLKAACIHSQWTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPP 592
DB 60 ISLMKDQVDQLQANGVAACLNSTQTRQQLEVMTCGTGQIRLLYIAPERLMLDNFLEH 119

QY 593 AAQLPPVAFACIDEAHCLSQWSHNRPCYLKRVCKLRRMGVHCFGLGTATATRTASDV 652
DB 120 LAHWNPVLLA-VDEAHCSQWGHDFRPEYALGQ-LRQRPPTLPFMALTATADTTTQDI 177

QY 653 AQHLVAEEDPLHGPAVPPTNLHLSVMSDRDTPQALLTLQKGFQNLDSI----- 703
DB 178 VRLLG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKFPLDQLMRVYVQSGR 221

QY 704 ---IYCNRRDTERIAALLRTCLHAAWVPGSGRAPKTTAEAYHAGMCSRERRVORAF 760
DB 222 KSGIYCNRAKVEDTAAAL-----QSKGISAAAYHAGLENNVRADVQEKF 267

QY 761 MQQLRVVATVAFGMGLDRPDVRAVLHLGLPPSFESYVQAVGRAGRDGQPAHCHFL-- 818
DB 268 QRDDLQIVVATVAFGMGINKPNRVFVHFDPINIESYVQETGRAGRDGLPAEAMLFYDP 327

QY 819 -----QPGEDLRELRRH 831
DB 328 ADMAWLRRCLEEKPGQ-LQDIERH 351

RESULT 9
US-09-753-143-77
; Sequence 77, Application US/09753143
; Patent No. 6838240
; GENERAL INFORMATION:
; APPLICANT: NATHAN A. ELLIS, JAMES GERMAN, AND JOANNA GRODEN
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF BLOOM'S SYNDROME
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: AMSTER, ROTHSTEIN & EBENSTEIN
; STREET: 90 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/753,143
; FILING DATE: 02-Jan-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/175,828
; FILING DATE: 1998-10-20
; ATTORNEY/AGENT INFORMATION:
; NAME: ELIZABETH A. BOGOSIAN
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/65
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: <Unknown>
; DESCRIPTION: OTHER NUCLEIC ACID
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
;
US-09-753-143-77

Query Match          9.7%; Score 621; DB 2; Length 370;
Best Local Similarity 40.0%; Pred. No. 7.1e-43;
Matches 154; Conservative 57; Mismatches 110; Indels 64; Gaps 12;

QY 474 EQLGHOAFRPGQRAVNRILSGISTLLVPTGAGKSLCYQLPALLYRRSPCLTLVVSPL 533
DB 4 ETFGYQFRPGQEEIITVLSGRDCLVVMPTGGKSLCYQIPALLNG----LTVVVSPL 59

QY 534 LSLMDDQVSGLPP-CLKAACIHSQWTRKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPP 592
DB 60 ISLMKDQVDQLQANGVAACLNSTQTRQQLEVMTCGTGQIRLLYIAPERLMLDNFLEH 119

QY 593 AAQLPPVAFACIDEAHCLSQWSHNRPCYLKRVCKLRRMGVHCFGLGTATATRTASDV 652
DB 120 LAHWNPVLLA-VDEAHCSQWGHDFRPEYALGQ-LRQRPPTLPFMALTATADTTTQDI 177

QY 653 AQHLVAEEDPLHGPAVPPTNLHLSVMSDRDTPQALLTLQKGFQNLDSI----- 703
DB 178 VRLLG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKFPLDQLMRVYVQSGR 221

QY 704 ---IYCNRRDTERIAALLRTCLHAAWVPGSGRAPKTTAEAYHAGMCSRERRVORAF 760
DB 222 KSGIYCNRAKVEDTAAAL-----QSKGISAAAYHAGLENNVRADVQEKF 267

QY 761 MQQLRVVATVAFGMGLDRPDVRAVLHLGLPPSFESYVQAVGRAGRDGQPAHCHFL-- 818
DB 268 QRDDLQIVVATVAFGMGINKPNRVFVHFDPINIESYVQETGRAGRDGLPAEAMLFYDP 327

QY 819 -----QPGEDLRELRRH 831
DB 328 ADMAWLRRCLEEKPGQ-LQDIERH 351

RESULT 10
US-09-463-702A-38
; Sequence 38, Application US/09463702A
; Patent No. 6335435
; GENERAL INFORMATION:
; APPLICANT: AGENE Research Institute, Co., Ltd.
; APPLICANT: HIRAKI AND ASSOCIATES
; APPLICANT: SHIMAMOTO, AKIRO
; APPLICANT: KITAO, SAORI
```

; APPLICANT: FURUICHI, YASUHIRO
; TITLE OF INVENTION: HUMAN GENE RECO4 ENCODING HELICASE
; FILE REFERENCE: HIRAI150
; CURRENT APPLICATION NUMBER: US/09/463,702A
; CURRENT FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/JP98/03114
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: JAPAN 9/200387
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-463-702A-38

Query Match 9.6%; Score 613.5; DB 2; Length 334;
Best Local Similarity 41.4%; Pred. No. 2.6e-42;
Matches 152; Conservative 53; Mismatches 111; Indels 51; Gaps 11;
QY 477 GHOAFRPGQERAVMRILSGISTLLVPTGAGKSLCYQLPALLYRRSPCLTLVVSPLLSL 536
Db 2 GYQFRPGQEBIIDTVLSGRDCLVVMPTGGKSLCYQIPALLNG---LTVVVSPLLSL 57
QY 537 MDDQVSGLPP-CLKAACIHSMTKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPPAAQ 595
Db 58 MKDQVDQLQANGVAAACLNSTQREQLLEVMTGCTGQIRLLYIAPERLMLDNFLEHLAH 117
QY 596 LPPVAFACIDEAHCLSQSHNFRPCYLVRCKVLRERMGVHCFLGTATATRTASDVAAQ 655
Db 118 WNPVLLA-VDEAHICISQWGHDFREYAAALGQ-LRQRPFLPFMALTAADTTTQDVR 175
QY 656 LAVAEBPLHGPAVPVPTNLHLSVSMRDQDQALLTLQGRFQNLDSI----- 703
Db 176 LG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKFPLDQLMRVYQEQKSG 219
QY 704 IYCNREDTERIAALLRTCLHAAPVPGSGRAPKTTAAAYHAGMCSRERRRVQAFMQG 763
Db 220 IYCNRAKVDTAARL-----OSKGISAAAYHAGLENNVRADVQEKFOR 265
QY 764 QLRVVVATVAFGMGLDRPDVAVLHGLPPSFESVVOAVGRAGDQOPAHCHLFLQPOGE 823
Db 266 DLQIVVATVAFGMGINKENVRVVFHFDIPRNIESYQETGRAGDGLPAEAMLFYDP--A 323
QY 824 DLRELRR 830
Db 324 DWAALRR 330

RESULT 11
US-09-699-135-38
; Sequence 38, Application US/09699135
; Patent No. 6472513
; GENERAL INFORMATION:
; APPLICANT: AGENE Research Institute, Co., Ltd.
; APPLICANT: HIRAKI AND ASSOCIATES
; APPLICANT: SHIMAMOTO, AKIRO
; APPLICANT: KITAO, SAORI
; APPLICANT: FURUICHI, YASUHIRO
; TITLE OF INVENTION: HUMAN GENE RECO4 ENCODING HELICASE
; FILE REFERENCE: HIRAI150
; CURRENT APPLICATION NUMBER: US/09/699,135
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US/09/463,702A
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/JP98/03114
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: JAPAN 9/200387
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38

; LENGTH: 334
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-699-135-38
Query Match 9.6%; Score 613.5; DB 2; Length 334;
Best Local Similarity 41.4%; Pred. No. 2.6e-42;
Matches 152; Conservative 53; Mismatches 111; Indels 51; Gaps 11;
QY 477 GHOAFRPGQERAVMRILSGISTLLVPTGAGKSLCYQLPALLYRRSPCLTLVVSPLLSL 536
Db 2 GYQFRPGQEBIIDTVLSGRDCLVVMPTGGKSLCYQIPALLNG---LTVVVSPLLSL 57
QY 537 MDDQVSGLPP-CLKAACIHSMTKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPPAAQ 595
Db 58 MKDQVDQLQANGVAAACLNSTQREQLLEVMTGCTGQIRLLYIAPERLMLDNFLEHLAH 117
QY 596 LPPVAFACIDEAHCLSQSHNFRPCYLVRCKVLRERMGVHCFLGTATATRTASDVAAQ 655
Db 118 WNPVLLA-VDEAHICISQWGHDFREYAAALGQ-LRQRPFLPFMALTAADTTTQDVR 175
QY 656 LAVAEBPLHGPAVPVPTNLHLSVSMRDQDQALLTLQGRFQNLDSI----- 703
Db 176 LG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKFPLDQLMRVYQEQKSG 219
QY 704 IYCNREDTERIAALLRTCLHAAPVPGSGRAPKTTAAAYHAGMCSRERRRVQAFMQG 763
Db 220 IYCNRAKVDTAARL-----OSKGISAAAYHAGLENNVRADVQEKFOR 265
QY 764 QLRVVVATVAFGMGLDRPDVAVLHGLPPSFESVVOAVGRAGDQOPAHCHLFLQPOGE 823
Db 266 DLQIVVATVAFGMGINKENVRVVFHFDIPRNIESYQETGRAGDGLPAEAMLFYDP--A 323
QY 824 DLRELRR 830
Db 324 DWAALRR 330
RESULT 12
US-09-489-039A-8075
; Sequence 8075, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8075
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8075
Query Match 9.4%; Score 606; DB 2; Length 619;
Best Local Similarity 35.5%; Pred. No. 2.7e-41;
Matches 169; Conservative 70; Mismatches 157; Indels 80; Gaps 18;
QY 474 EOLGHQAFRPGQERAVMRILSGISTLLVPTGAGKSLCYQLPALLYRRSPCLTLVVSPL 533
Db 30 ETFGYQFRPGQETIETALEGRDCLVVMPTGGKSLCYQVPALVMGG---LTVVVSPL 85
QY 534 LSLMDDQVSG-LPPCLKAACIHSMTKQRESVLQKIRAAQVHVLMLTPEALVGAGGLPP 592
Db 86 ISLMKDDQVDQLANGVAAACLNSTQREQQEQWAGCSGQVRLLYIAPERLMLDNFLEH 145
QY 593 AAQLPPVAFACIDEAHCLSQSHNFRPCYLVRCKVLRERMGVHCFLGTATATRTASDV 652
Db 146 LANW-NLAWLAVDEAHICISQWGHDFREYAAALGQ-LRQMPQIPFMALTAADTTTTRDI 203

Db 637 VNRARTPASAAGT-----QPERVRRAPTAPALRPLK----- 671
Qy 473 LEQLGHQAFRQGERAVNRILSGISLLVLPFGAGKSLCYQLPALLYRRSPFCLTLVVSP 532
Db 672 -EAFGSDFRFVQEAACVACRAATAGEDLLVMPGTGAGKSLCYQLPGLARAG----TTLVVSP 726
Qy 533 LLSLMDQVSGLPPC-LXAACIHSGMTKQRESVLQKTRAAQVHVLMLTPEALVAGAGLGP 591
Db 727 LIALMEDQVRLQSLGFAADRIHSGRDRAMSRQVCADYLEDRDLDFLTAPERLGVPGFVE 786
Qy 592 PAALPPVAFACIDEACHLSQWSHNFPCYLKRVCKVLRRMGV---HCFLGLTATATRT 648
Db 787 FLARTP-ALTAVDAHCISQWCHDFRFDY----RLGARPLRLPAPVVALTATATPDV 841
Qy 649 ASDVAQHLAVAEEDLHGAPVP-----TNLHLSVSMRDRTD-----QALLTLQ 694
Db 842 QRDIVQQLG-----LQFGGKARTFIHGFRRTNIAIEV---RELNPGARGDAIQLLED 892
Qy 695 KRFQNLDSIIICNRRDTERIAALLRTCLHAAWVPGSGGRAPKTTAEAYHAGMCSRERR 754
Db 893 EENR---PAIVYAATRKHAQLADQL-----AGEFP---AAAYHAGLQPSERD 934
Qy 755 RVORAFMOGLRVVATVAFMGDLDPDVRAVLHLGLPPSPESYVQAVGRAGRDGQPAHC 814
Db 935 RVQAEFLRGSLEVIATVAFMGIDKADVTVIHAALPASLEGYYQELGRAGRDGKPSRA 994
Qy 815 ---HLFLOQGEDLRELRRHVHADSTDFLAVKRLVQRFVFPACTCTCTRPSPSEQEGAVGE 871
Db 995 VLLHSYIDRRTHFPFRHDYPEA-----YVLERLFKS-TAPQLRPFKAVLQGRVRGD 1044
Qy 872 RPVKYPPQEAQOL 885
Db 1045 ---PEVFDKALEQL 1055

RESULT 15

US-09-134-000C-5710
; Sequence 5710, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5710
; LENGTH: 589
; TYPE: PR1
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5710

Query Match 8.6%; Score 554.5; DB 2; Length 589;
Best Local Similarity 34.0%; Pred. No. 4.6e-37;
Matches 143; Conservative 67; Mismatches 147; Indels 63; Gaps 10;

Qy 472 ALEQL-----GHAQFRPQGERAVMRILSGISLLVLPFGAGKSLCYQLPALLYRRSPCL 526
Db 3 ALQELLKDTFGYDDFRPGQETIIRHVLRQENVLGMPTGGKSIQVLPALLDN----L 58
Qy 527 TLVSPSLSLMDQVS-----GLPCLKAACHSGMTRKQRESVLQKTRAAQVHVLMLTLP 581
Db 59 TLVISPLISLMDQVNDALNMGIP-----ATYINSTISYQEMNHRIQLAVNKEVKLLYVAP 114
Qy 582 EALVGAGGLPPAAQLPPVAFACIDEACHLSQWSHNFPCYLKRVCKVLRRMGVHCFGLT 641
Db 115 ERLESTDFQQLMTHV-PIDLLAVDEAHACISQWCHDFRPSYLRARLAIIDFQQQPTVIALT 173

Qy 642 ATATRTASDVAQHLAVAEEDLHGAPVPPTNLHLSVSMRDRTDQALLTLTLLQGRKFONLD 701
Db 174 ATATPQVAEDIVKQLRIPSENEIK-TGFARENLSFQVVKQDNRDVFLLEYLKNMTGQ--- 229
Qy 702 SIIICNRRDTERIAALLRTCLHAAWVPGSGGRAPKTTAEAYHAGMCSRERRRVQRAFM 761
Db 230 SGIIVASTRKEVERIVHLL-----ESKKIAAGMYHGGMSEQLRSENQEAFL 275
Qy 762 QGQLRVVATVATGCMGLDRPDVRAVLHLGLPPSPESYVQAVGRAGRDGQPAHCHLFLQ 821
Db 276 YDQVQVMVATNAFGMGINKSNVRVTHAQVFGNIESYYQAGRAGRDGLPSDAVLMFAPQ 335
Qy 822 GED-----LRELRRHVHADSTDFLAVKRLVQRFVFPACTCTCTR 859
Db 336 DLQIQYFIEQSENTIDYKQKEYLKLREMSQYANAQ-----MCLQKYLIRYFGEESTDCGR 391

Search completed: December 27, 2005, 21:55:16
Job time : 55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:53:22 ; Search time 178.4 Seconds
(without alignments)
2829.248 Million cell updates/sec

Title: US-09-889-325-4

Perfect score: 6424

Sequence: 1 MERLDVRELRQAWERAFRR.....HLSPHALVGLATELLQVAR 1208

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|----------------------|
| 1 | 6424 | 100.0 | 1208 | 3 | US-09-889-325-4 |
| 2 | 6424 | 100.0 | 1208 | 5 | US-10-723-860-4382 |
| 3 | 6424 | 100.0 | 1208 | 5 | US-10-882-403-4 |
| 4 | 6424 | 100.0 | 1208 | 5 | US-10-756-149-5768 |
| 5 | 3944 | 61.4 | 1216 | 6 | US-11-084-955-2 |
| 6 | 1723.5 | 26.8 | 1579 | 6 | US-11-097-143-16743 |
| 7 | 765 | 11.9 | 874 | 4 | US-10-437-963-121617 |
| 8 | 632.5 | 9.8 | 607 | 4 | US-10-374-077-75 |
| 9 | 631.5 | 9.8 | 610 | 3 | US-09-815-242-10414 |
| 10 | 631.5 | 9.8 | 610 | 4 | US-10-282-122A-43357 |
| 11 | 621 | 9.7 | 370 | 3 | US-09-753-143-77 |
| 12 | 619 | 9.6 | 609 | 4 | US-10-282-122A-56210 |
| 13 | 609.5 | 9.5 | 609 | 4 | US-10-282-122A-75783 |
| 14 | 609.5 | 9.5 | 615 | 3 | US-09-815-242-13747 |
| 15 | 607.5 | 9.5 | 615 | 4 | US-10-282-122A-74973 |
| 16 | 606 | 9.4 | 608 | 4 | US-10-282-122A-59929 |
| 17 | 593 | 9.2 | 610 | 4 | US-10-282-122A-78025 |
| 18 | 589.5 | 9.2 | 600 | 4 | US-10-282-122A-69214 |
| 19 | 581.5 | 9.1 | 620 | 4 | US-10-282-122A-76978 |
| 20 | 578.5 | 9.0 | 632 | 4 | US-10-282-122A-67293 |
| 21 | 578 | 9.0 | 619 | 3 | US-09-815-242-11085 |
| 22 | 578 | 9.0 | 619 | 4 | US-10-282-122A-58271 |
| 23 | 554.5 | 8.6 | 589 | 4 | US-10-282-122A-42447 |
| 24 | 552 | 8.6 | 590 | 4 | US-10-282-122A-60691 |
| 25 | 549 | 8.5 | 597 | 4 | US-10-282-122A-72864 |
| 26 | 547.5 | 8.5 | 589 | 3 | US-09-815-242-10803 |
| 27 | 545 | 8.5 | 580 | 3 | US-09-815-242-4959 |

ALIGNMENTS

RESULT 1

US-09-889-325-4
; Sequence 4, Application US/09889325
; Publication No. US20040224312A1
; GENERAL INFORMATION:

; APPLICANT: AGENE Research Institute, Co., Ltd.
; TITLE OF INVENTION: Gene causative of Rothmund-Thomson syndrome
; TITLE OF INVENTION: and its gene product
; FILE REFERENCE: A1-003PCT
; CURRENT APPLICATION NUMBER: US/09/889,325
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: JP 1999-11218
; PRIOR FILING DATE: 1999-01-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-889-325-4

Query Match 100.0%; Score 6424; DB 3; Length 1208;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| Qy | 1 | MERLDVRELRQAWERAFRRQRRPSODDVEAAPEETRALYREYRTLKRTTQAGGGLR | 60 |
| Db | 1 | MERLDVRELRQAWERAFRRQRRPSODDVEAAPEETRALYREYRTLKRTTQAGGGLR | 60 |
| Qy | 61 | SSESLPAAAEAEAPRCMGPHLNRAATKSPOPTGSRSGSVDPYQGRLLKANLKTLOAG | 120 |
| Db | 61 | SSESLPAAAEAEAPRCMGPHLNRAATKSPOPTGSRSGSVDPYQGRLLKANLKTLOAG | 120 |
| Qy | 121 | PALGRPRPPLGRASSKASTPKPGTGPVPSFAEKVSDPPQLPEPPQPRGLQHLQASIS | 180 |
| Db | 121 | PALGRPRPPLGRASSKASTPKPGTGPVPSFAEKVSDPPQLPEPPQPRGLQHLQASIS | 180 |
| Qy | 181 | QRLGSLDPGWLQCHSEYVDFLGAPKACRPDLGSEESOLLIPGESAVLPGAGSGGPEAS | 240 |
| Db | 181 | QRLGSLDPGWLQCHSEYVDFLGAPKACRPDLGSEESOLLIPGESAVLPGAGSGGPEAS | 240 |
| Qy | 241 | AFQEVSIKVGSPQSSSGCKRRWNEEPWESPAQVQOQSSQAGPPSEGAGAVAEEDPPG | 300 |
| Db | 241 | AFQEVSIKVGSPQSSSGCKRRWNEEPWESPAQVQOQSSQAGPPSEGAGAVAEEDPPG | 300 |
| Qy | 301 | EPVQAQPPQPCSSPSNPRYHGLSPSSQARAGKAEGTAPLHIFPRLARHDRGNVRLNMKQ | 360 |
| Db | 301 | EPVQAQPPQPCSSPSNPRYHGLSPSSQARAGKAEGTAPLHIFPRLARHDRGNVRLNMKQ | 360 |
| Qy | 361 | KHYVGRALRSLLRKQAWKQWRKKGECFGGGGATVTTKSCFLNEQPDHWAACQPRPA | 420 |

Sequence 11915, A
Sequence 66476, A
Sequence 64441, A
Sequence 8990, Ap
Sequence 3, Appli
Sequence 141345,
Sequence 57618, A
Sequence 70676, A
Sequence 7447, Ap
Sequence 76, Appl
Sequence 12385, A
Sequence 256333,
Sequence 67521, A
Sequence 59459, A
Sequence 53151, A
Sequence 70264, A
Sequence 12306, A
Sequence 48328, A

28 545 8.5 712 3 US-09-815-242-11915
29 545 8.5 712 4 US-10-282-122A-66476
30 544.5 8.5 390 4 US-10-282-122A-61441
31 544 8.5 341 4 US-10-369-493-8990
32 542.5 8.4 1955 4 US-10-293-504-3
33 541 8.4 588 4 US-10-437-963-141345
34 528 8.2 590 4 US-10-282-122A-57618
35 526 8.2 592 4 US-10-282-122A-70676
36 526 8.2 637 4 US-10-724-972A-7447
37 525.5 8.2 1328 4 US-10-374-077-76
38 524.5 8.2 658 4 US-10-156-761-12385
39 519.5 8.1 618 4 US-10-425-115-256333
40 518.5 8.1 715 4 US-10-282-122A-67521
41 517 8.0 653 4 US-10-425-114-59459
42 515 8.0 827 4 US-10-282-122A-53151
43 514 8.0 593 4 US-10-282-122A-70264
44 513 8.0 593 3 US-09-815-242-12306
45 511 8.0 601 4 US-10-282-122A-48328

Db 361 KHYVGRALRRLRLKQAWKQKWKGCFCGGGATVTTKSCFLNEQFDHWAACQPRPA 420
Qy 421 SEEDTDVAGPEPLVPSPQVPEVPSLDPTVLPLSLGSPGQLAETPAEVFQALQOLGHOA 480
Db 421 SEEDTDVAGPEPLVPSPQVPEVPSLDPTVLPLSLGSPGQLAETPAEVFQALQOLGHOA 480
Qy 481 FRPQGERAVMRILSGISTLLVPTGAGKSLCYQLPALLYSRSPCLTLVWSPLLSLMDDQ 540
Db 481 FRPQGERAVMRILSGISTLLVPTGAGKSLCYQLPALLYSRSPCLTLVWSPLLSLMDDQ 540
Qy 541 VSGLPCLKAACIHSWTRKQRESVLQKIRAAQVHVLMLTPEALVAGAGLPPAAQLPPVA 600
Db 541 VSGLPCLKAACIHSWTRKQRESVLQKIRAAQVHVLMLTPEALVAGAGLPPAAQLPPVA 600
Qy 601 FACIDEAHCLSQWSHNFRPCYLKVRCKLVRERMGVHCFGLGTATATRTASDVVAQHLAAVE 660
Db 601 FACIDEAHCLSQWSHNFRPCYLKVRCKLVRERMGVHCFGLGTATATRTASDVVAQHLAAVE 660
Qy 661 EPDLHGPAVPVPTNLHLVSMDRDTDQALLTLLOGKRFQNLDSIIYCNRRREDTERIAALL 720
Db 661 EPDLHGPAVPVPTNLHLVSMDRDTDQALLTLLOGKRFQNLDSIIYCNRRREDTERIAALL 720
Qy 721 RTCLHAAWPGSGGRAPKTTAEAYHAGMCSRERRRVRQAFMQQLRVVVVATVAFGMGLDR 780
Db 721 RTCLHAAWPGSGGRAPKTTAEAYHAGMCSRERRRVRQAFMQQLRVVVVATVAFGMGLDR 780
Qy 781 PDVRAVLHLGLPPSPFESYVQAVGRAGDQGAHCHLFLQPGEDLRELRRHVHADSTDFL 840
Db 781 PDVRAVLHLGLPPSPFESYVQAVGRAGDQGAHCHLFLQPGEDLRELRRHVHADSTDFL 840
Qy 841 AVKRLVQVFPACTCTCTRPPSEQEGAVGGERPVPKYPQAEQLSHQAAPGPRVCMGH 900
Db 841 AVKRLVQVFPACTCTCTRPPSEQEGAVGGERPVPKYPQAEQLSHQAAPGPRVCMGH 900
Qy 901 ERALPIQLTVQALDMPAEIATILCYLELHPHMLELLATYTHCRNLNCPGPAQLQALA 960
Db 901 ERALPIQLTVQALDMPAEIATILCYLELHPHMLELLATYTHCRNLNCPGPAQLQALA 960
Qy 961 HRCPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGV 1020
Db 961 HRCPLAVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGV 1020
Qy 1021 RRGTVLVFSELAFLHRLSPGDLTAEKQDQICDFLYGRVQARERQALRLRRTQAFHSV 1080
Db 1021 RRGTVLVFSELAFLHRLSPGDLTAEKQDQICDFLYGRVQARERQALRLRRTQAFHSV 1080
Qy 1081 APPSCGCGLEQDDEERSTRLLKDLGRYFEEBEGQEPGEMEDAQGPFGQARLQWEDQVR 1140
Db 1081 APPSCGCGLEQDDEERSTRLLKDLGRYFEEBEGQEPGEMEDAQGPFGQARLQWEDQVR 1140
Qy 1141 CDIRQFLSLRPEEKFSFRAVARIIPHGIGSPCYPAQVYQDQRRFWRKYLHLHSFHALVGLAT 1200
Db 1141 CDIRQFLSLRPEEKFSFRAVARIIPHGIGSPCYPAQVYQDQRRFWRKYLHLHSFHALVGLAT 1200
Qy 1201 EELLQVAR 1208
Db 1201 EELLQVAR 1208

RESULT 2
US-10-723-860-4382
; Sequence 4382, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26

; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4382
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4382

Query Match 100.0%; Score 6424; DB 5; Length 1208;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MERLRDVRERLQAWERAFRRQRRRSQDDVEAARPEETRALYREYRTLKRTTGQAGGLR 60
Db 1 MERLRDVRERLQAWERAFRRQRRRSQDDVEAARPEETRALYREYRTLKRTTGQAGGLR 60
Qy 61 SSESPPAAAEAEPEPCWGPFLNRAATKSPQTPGRSRQGSVPDYGQRLKANLKGTLQAG 120
Db 61 SSESPPAAAEAEPEPCWGPFLNRAATKSPQTPGRSRQGSVPDYGQRLKANLKGTLQAG 120
Qy 121 PALGRRPWLGRASSKASTPKPGTGPVPSFAEKVSDPEPPQPPORPORLOHLOASLS 180
Db 121 PALGRRPWLGRASSKASTPKPGTGPVPSFAEKVSDPEPPQPPORPORLOHLOASLS 180
Qy 181 QRLGSLDPGWLQCHSEVPDFLGAAPKACRPDLGSEESQLLIPGESAVLGPAGCSQGPAS 240
Db 181 QRLGSLDPGWLQCHSEVPDFLGAAPKACRPDLGSEESQLLIPGESAVLGPAGCSQGPAS 240
Qy 241 AFEVSIIRVGSPPSSSGGKREWNEEPWESPAQVQOESSQAGPPPSSEGAGAVAEEDPPG 300
Db 241 AFEVSIIRVGSPPSSSGGKREWNEEPWESPAQVQOESSQAGPPPSSEGAGAVAEEDPPG 300
Qy 301 EPVQAPPQPCSSPNRPHGLSPSSQARAGKAEGTAPLHIFPRLARHDRGNVYRLNNKQ 360
Db 301 EPVQAPPQPCSSPNRPHGLSPSSQARAGKAEGTAPLHIFPRLARHDRGNVYRLNNKQ 360
Qy 361 KHYVGRALRSLRLKQAWKQKWKGCFCGGGATVTTKSCFLNEQFDHWAACQPRPA 420
Db 361 KHYVGRALRSLRLKQAWKQKWKGCFCGGGATVTTKSCFLNEQFDHWAACQPRPA 420
Qy 421 SEEDTDVAGPEPLVPSPQVPEVPSLDPTVLPLSLGSPGQLAETPAEVFQALQOLGHOA 480
Db 421 SEEDTDVAGPEPLVPSPQVPEVPSLDPTVLPLSLGSPGQLAETPAEVFQALQOLGHOA 480
Qy 481 FRPQGERAVMRILSGISTLLVPTGAGKSLCYQLPALLYSRSPCLTLVWSPLLSLMDDQ 540
Db 481 FRPQGERAVMRILSGISTLLVPTGAGKSLCYQLPALLYSRSPCLTLVWSPLLSLMDDQ 540
Qy 541 VSGLPCLKAACIHSWTRKQRESVLQKIRAAQVHVLMLTPEALVAGAGLPPAAQLPPVA 600
Db 541 VSGLPCLKAACIHSWTRKQRESVLQKIRAAQVHVLMLTPEALVAGAGLPPAAQLPPVA 600
Qy 601 FACIDEAHCLSQWSHNFRPCYLKVRCKLVRERMGVHCFGLGTATATRTASDVVAQHLAAVE 660
Db 601 FACIDEAHCLSQWSHNFRPCYLKVRCKLVRERMGVHCFGLGTATATRTASDVVAQHLAAVE 660
Qy 661 EPDLHGPAVPVPTNLHLVSMDRDTDQALLTLLOGKRFQNLDSIIYCNRRREDTERIAALL 720
Db 661 EPDLHGPAVPVPTNLHLVSMDRDTDQALLTLLOGKRFQNLDSIIYCNRRREDTERIAALL 720
Qy 721 RTCLHAAWPGSGGRAPKTTAEAYHAGMCSRERRRVRQAFMQQLRVVVVATVAFGMGLDR 780
Db 721 RTCLHAAWPGSGGRAPKTTAEAYHAGMCSRERRRVRQAFMQQLRVVVVATVAFGMGLDR 780
Qy 781 PDVRAVLHLGLPPSPFESYVQAVGRAGDQGAHCHLFLQPGEDLRELRRHVHADSTDFL 840
Db 781 PDVRAVLHLGLPPSPFESYVQAVGRAGDQGAHCHLFLQPGEDLRELRRHVHADSTDFL 840
Qy 841 AVKRLVQVFPACTCTCTRPPSEQEGAVGGERPVPKYPQAEQLSHQAAPGPRVCMGH 900
Db 841 AVKRLVQVFPACTCTCTRPPSEQEGAVGGERPVPKYPQAEQLSHQAAPGPRVCMGH 900

QY 901 ERALPIQLTVQALDMPERAIETLLCYLELHPHMLLELLATTYTHCRNLCPGPGPAQLQALA 960
 Db 901 ERALPIQLTVQALDMPERAIETLLCYLELHPHMLLELLATTYTHCRNLCPGPGPAQLQALA 960
 QY 961 HRCPPPLAVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRTGV 1020
 Db 961 HRCPPPLAVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRTGV 1020
 QY 1021 RRGTVLVFSELAFLHRS PGDLTAEEKDQICDFLYGRVQARERQALRLRTTFOAFHSV 1080
 Db 1021 RRGTVLVFSELAFLHRS PGDLTAEEKDQICDFLYGRVQARERQALRLRTTFOAFHSV 1080
 QY 1081 APPSCGCPLEQDDEERSTRLLKDLGRYFEEREGQPGMGMEDAQGPQOARLQDWDQVR 1140
 Db 1081 APPSCGCPLEQDDEERSTRLLKDLGRYFEEREGQPGMGMEDAQGPQOARLQDWDQVR 1140
 QY 1141 CDIROFLSLRPEEKFSRAVARIFHGIGSPCYPAQVYGQDRFRWKYHLHLSPHALVGLAT 1200
 Db 1141 CDIROFLSLRPEEKFSRAVARIFHGIGSPCYPAQVYGQDRFRWKYHLHLSPHALVGLAT 1200
 QY 1201 BELLQVAR 1208
 Db 1201 BELLQVAR 1208
 RESULT 3
 US-10-882-405-4
 ; Sequence 4, Application US/10882405
 ; Publication No. US20040259147A1
 ; GENERAL INFORMATION:
 ; APPLICANT: AGENE Research Institute, Co., Ltd.
 ; TITLE OF INVENTION: Gene causative of Rothmund-Thomson syndrome
 ; TITLE OF INVENTION: and its gene product
 ; FILE REFERENCE: A1-003PCT
 ; CURRENT APPLICATION NUMBER: US/10/882,405
 ; CURRENT FILING DATE: 2004-07-02
 ; PRIOR APPLICATION NUMBER: US/09/889,325
 ; PRIOR FILING DATE: 2001-12-31
 ; PRIOR APPLICATION NUMBER: JP 1999-11218
 ; PRIOR FILING DATE: 1999-01-19
 ; NUMBER OF SEQ ID NOS: 46
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 1208
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-882-405-4
 Query Match 100.0%; Score 6424; DB 5; Length 1208;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MERLDRVRELQWERAFAFRQRGRPSQDDVEAAPETRALRYEYTLKRTTQAGGGLR 60
 Db 1 MERLDRVRELQWERAFAFRQRGRPSQDDVEAAPETRALRYEYTLKRTTQAGGGLR 60
 QY 61 SSELPAABEAEPPRCWPHLNRAATKSPQTPGRSROGSVDYQORLKANLKGTLQAG 120
 Db 61 SSELPAABEAEPPRCWPHLNRAATKSPQTPGRSROGSVDYQORLKANLKGTLQAG 120
 QY 121 PALGREPWLGRASSKASTPKPGTGFVPSFAEKVSDPEPQLPEPQPRGRQLQASLS 180
 Db 121 PALGREPWLGRASSKASTPKPGTGFVPSFAEKVSDPEPQLPEPQPRGRQLQASLS 180
 QY 181 QRLGSLDPGWLQCHSEVPDPLGAPKACRPDLGSEESQLLIPGESAVLPGFAGSQGPAS 240
 Db 181 QRLGSLDPGWLQCHSEVPDPLGAPKACRPDLGSEESQLLIPGESAVLPGFAGSQGPAS 240
 QY 241 AFOEVSIRVGSPOPSSGGGKRWNEEPWESPAQVOQESSQAGPPSEGAGAVBEDPPG 300
 Db 241 AFOEVSIRVGSPOPSSGGGKRWNEEPWESPAQVOQESSQAGPPSEGAGAVBEDPPG 300

QY 301 EPVQAQPPQCSPPSNRYHGLSPSSQARAKAGTAPLHIFPRLARHNRGVRLNWKQ 360
 Db 301 EPVQAQPPQCSPPSNRYHGLSPSSQARAKAGTAPLHIFPRLARHNRGVRLNWKQ 360
 QY 361 KHYVYRGALRSRLRKQAKWKKEGCEFGGGATVTTKESCFNLNQFQDHAAQCPRA 420
 Db 361 KHYVYRGALRSRLRKQAKWKKEGCEFGGGATVTTKESCFNLNQFQDHAAQCPRA 420
 QY 421 SEEDTAVGPEPLVPSQPPVEPSLPTVLPYLSIGSPSQLAETPAEVAOLQOLGHOA 480
 Db 421 SEEDTAVGPEPLVPSQPPVEPSLPTVLPYLSIGSPSQLAETPAEVAOLQOLGHOA 480
 QY 481 FRPQERAVMRLSGISTLLVLPAGKSLCYQLPALLYSRSSPCLTLVSPFLSLMDQ 540
 Db 481 FRPQERAVMRLSGISTLLVLPAGKSLCYQLPALLYSRSSPCLTLVSPFLSLMDQ 540
 QY 541 VSGLPPLCKAACTHSGMTRKQRESVLQIRAAQVHVLMLTPEALVGAGGLPPAAQLPPVA 600
 Db 541 VSGLPPLCKAACTHSGMTRKQRESVLQIRAAQVHVLMLTPEALVGAGGLPPAAQLPPVA 600
 QY 601 PACIDEAHCLSQWSHNFRPCYLKVRKREMGVHCFLGTATATRTTASDVAQHLVAE 660
 Db 601 PACIDEAHCLSQWSHNFRPCYLKVRKREMGVHCFLGTATATRTTASDVAQHLVAE 660
 QY 661 EPDLHGPAVPNTLHLSVSMRDRTDQALLTLQGRFONLDSIIYCNRRDTERIAALL 720
 Db 661 EPDLHGPAVPNTLHLSVSMRDRTDQALLTLQGRFONLDSIIYCNRRDTERIAALL 720
 QY 721 RTCLHAAMVPGSGGRAPKTTAEAYHAGMCSRERRRVQAFMQGQLRVVVVATVAFMGGLDR 780
 Db 721 RTCLHAAMVPGSGGRAPKTTAEAYHAGMCSRERRRVQAFMQGQLRVVVVATVAFMGGLDR 780
 QY 781 PDVRAVLHGLPPSPFESTVQVAGRAGRDQPAHCHLFLQPGQEDLRELRRHVHADSTDFL 840
 Db 781 PDVRAVLHGLPPSPFESTVQVAGRAGRDQPAHCHLFLQPGQEDLRELRRHVHADSTDFL 840
 QY 841 AVKRLVQRFVFACTCTCTRPSEOBGAVGGERPVKYPPOEAEQLSHQAAQPGPRVCMGH 900
 Db 841 AVKRLVQRFVFACTCTCTRPSEOBGAVGGERPVKYPPOEAEQLSHQAAQPGPRVCMGH 900
 QY 901 ERALPIQLTVQALDMPERAIETLLCYLELHPHMLLELLATTYTHCRNLCPGPGPAQLQALA 960
 Db 901 ERALPIQLTVQALDMPERAIETLLCYLELHPHMLLELLATTYTHCRNLCPGPGPAQLQALA 960
 QY 961 HRCPPPLAVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRTGV 1020
 Db 961 HRCPPPLAVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRTGV 1020
 QY 1021 RRGTVLVFSELAFLHRS PGDLTAEEKDQICDFLYGRVQARERQALRLRTTFOAFHSV 1080
 Db 1021 RRGTVLVFSELAFLHRS PGDLTAEEKDQICDFLYGRVQARERQALRLRTTFOAFHSV 1080
 QY 1081 APPSCGCPLEQDDEERSTRLLKDLGRYFEEREGQPGMGMEDAQGPQOARLQDWDQVR 1140
 Db 1081 APPSCGCPLEQDDEERSTRLLKDLGRYFEEREGQPGMGMEDAQGPQOARLQDWDQVR 1140
 QY 1141 CDIROFLSLRPEEKFSRAVARIFHGIGSPCYPAQVYGQDRFRWKYHLHLSPHALVGLAT 1200
 Db 1141 CDIROFLSLRPEEKFSRAVARIFHGIGSPCYPAQVYGQDRFRWKYHLHLSPHALVGLAT 1200
 QY 1201 BELLQVAR 1208
 Db 1201 BELLQVAR 1208

RESULT 4

US-10-756-149-5768
 ; Sequence 5768, Application US/10756149
 ; Publication No. US20050181375A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Narasba
 ; APPLICANT: Zlotnik, Albert
 ; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND

```
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5768
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5768

Query Match      100.0%; Score 6424; DB 5; Length 1208;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MERLRDVRERLQAWERAFRRQRGRPPSQDDVEAAPEETRALYRYRTLKRTTGAGGGLR 60
Db 1 MERLRDVRERLQAWERAFRRQRGRPPSQDDVEAAPEETRALYRYRTLKRTTGAGGGLR 60
Qy 61 SSESIPAAAEAPRRCWGHPLNRAATKSPQTPGRSRQSGVDPYQORLKANLKGTLQAG 120
Db 61 SSESIPAAAEAPRRCWGHPLNRAATKSPQTPGRSRQSGVDPYQORLKANLKGTLQAG 120
Qy 121 PALGRRPWPLGRASKASTPKPGTGPVPSFAEKVSDPEPPQPPRPGRLQHLQASLS 180
Db 121 PALGRRPWPLGRASKASTPKPGTGPVPSFAEKVSDPEPPQPPRPGRLQHLQASLS 180
Qy 181 ORLGSLDQGLWLRCHSEVDPFLGAPKACRPDLGSEESQLLIPGESAVILPGAGSQGPAS 240
Db 181 ORLGSLDQGLWLRCHSEVDPFLGAPKACRPDLGSEESQLLIPGESAVILPGAGSQGPAS 240
Qy 241 AFQVSVIRVSGPQPSGSGEKRRWNEPWPESPAQVQOESSQAGPPSEGAGAVAEEDPPG 300
Db 241 AFQVSVIRVSGPQPSGSGEKRRWNEPWPESPAQVQOESSQAGPPSEGAGAVAEEDPPG 300
Qy 301 EPVQAQPPQPCSSPNRYHGLSPSSQARAGKAGTAPLHIFPRLARHNRGNYVRLNMKQ 360
Db 301 EPVQAQPPQPCSSPNRYHGLSPSSQARAGKAGTAPLHIFPRLARHNRGNYVRLNMKQ 360
Qy 361 KHYVRGALRSRLRKQAWKQKWKCECFGGGATVTTKESCFLNQDFHWAACQRPRA 420
Db 361 KHYVRGALRSRLRKQAWKQKWKCECFGGGATVTTKESCFLNQDFHWAACQRPRA 420
Qy 421 SEEDTDVAGPEPLVPSQPVEVPSLPTVPLPLYSGLSPGQLAETPAEVFOALBOLGHOA 480
Db 421 SEEDTDVAGPEPLVPSQPVEVPSLPTVPLPLYSGLSPGQLAETPAEVFOALBOLGHOA 480
Qy 481 FRPQGERAVMRILSGISTLLVLTGAGKSLCYQLPALLYRRSPCLTLVWSPLLSLMDQ 540
Db 481 FRPQGERAVMRILSGISTLLVLTGAGKSLCYQLPALLYRRSPCLTLVWSPLLSLMDQ 540
Qy 541 VSGLPCCCKAACIHSGMTRKQRESVLQIRAAQVHVLMLTPEALVAGAGGLPPAAQLPPVA 600
Db 541 VSGLPCCCKAACIHSGMTRKQRESVLQIRAAQVHVLMLTPEALVAGAGGLPPAAQLPPVA 600
Qy 601 FACIDEAHCLSQWGHNRPFCVLRCKVLRERMGVHCFGLGTATATRTASDVAAHVAE 660
Db 601 FACIDEAHCLSQWGHNRPFCVLRCKVLRERMGVHCFGLGTATATRTASDVAAHVAE 660
Qy 661 EPDLHGPAVPTNLHLVSMRDRDQALLTLLOGKRFQNLDSIIICNRREDTERIAALL 720
Db 661 EPDLHGPAVPTNLHLVSMRDRDQALLTLLOGKRFQNLDSIIICNRREDTERIAALL 720
Qy 721 RTCLHAHVPSCGGRAPKTTAAEHAGCSERRRVRQAFMQGQLRVVVAATVAFGMGLDR 780
Db 721 RTCLHAHVPSCGGRAPKTTAAEHAGCSERRRVRQAFMQGQLRVVVAATVAFGMGLDR 780
Qy 781 PDVRAVLHLGLPPSFESVQAVGAGRDGQPAHCHFLQPOGEDRLRLRHVHADSTDFL 840
Db 781 PDVRAVLHLGLPPSFESVQAVGAGRDGQPAHCHFLQPOGEDRLRLRHVHADSTDFL 840
Qy 841 AVKRLVQRVFPACTCTCTRPSPSEOGAVGGERPVPKYPQOEALSHQAAAPGPRRVCWG 900
```

```
Db 841 AVKRLVQRVFPACTCTCTRPSPSEOGAVGGERPVPKYPQOEALSHQAAAPGPRRVCWG 900
Qy 901 ERALPIQLTVOALDMPPEAEITLLCYLELHPHWELELLATYTHCRINCPCGPAQLQALA 960
Db 901 ERALPIQLTVOALDMPPEAEITLLCYLELHPHWELELLATYTHCRINCPCGPAQLQALA 960
Qy 961 HRCPPPLAVCLAQQLPEDPGOGSSSVFDMVKLVDSMGWELASVRRALCOLQDWDHEPTGV 1020
Db 961 HRCPPPLAVCLAQQLPEDPGOGSSSVFDMVKLVDSMGWELASVRRALCOLQDWDHEPTGV 1020
Qy 1021 RRGTVLVFSELAFLHLSRSPDLTAAEKQOICDFLYGRVQARERQALRLRRTFOAFHSV 1080
Db 1021 RRGTVLVFSELAFLHLSRSPDLTAAEKQOICDFLYGRVQARERQALRLRRTFOAFHSV 1080
Qy 1081 AFPSCGPCLFQOQDEERSTRLKDLGLGRYFEFEEQSGPGMEDAQGPFGQARLQDWEQVR 1140
Db 1081 AFPSCGPCLFQOQDEERSTRLKDLGLGRYFEFEEQSGPGMEDAQGPFGQARLQDWEQVR 1140
Qy 1141 CDIRQFLSLRPEKFSRAVARIFPHGIGSPCYPAQVYGQDRRFWRKYLHLSFHALVGLAT 1200
Db 1141 CDIRQFLSLRPEKFSRAVARIFPHGIGSPCYPAQVYGQDRRFWRKYLHLSFHALVGLAT 1200
Qy 1201 EELLQVAR 1208
Db 1201 EELLQVAR 1208

RESULT 5
US-11-084-955-2
; Sequence 2, Application US/11084955
; Publication No. US20050183149A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Radiological Sciences
; TITLE OF INVENTION: Targeted knockout mouse for human Rothmund-Thomson syndrome and
; FILE REFERENCE: YIL-0610
; CURRENT APPLICATION NUMBER: US/11/084,955
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: JP 2003-185409
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1216
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-084-955-2

Query Match      61.4%; Score 3944; DB 6; Length 1216;
Best Local Similarity 63.9%; Pred. No. 1.2e-248;
Matches 795; Conservative 110; Mismatches 267; Indels 72; Gaps 13;

Qy 1 MERLRDVRERLQAWERAFRRQRGRPPSQDDVEAAPEETRALYRYRTLKRTTGAGGGLR 60
Db 1 MERLATVTRALQWERAFARLHGRPPAKGDVEAAPEETRALYREYNLQAVRQADDRHR 60
Qy 61 SSE-SLPAAABEAPRRCWGHPLNRAATKSPQTPGRSRQSGVDPYQORLKANLKGTLQ 119
Db 61 VLEQSLAEAAEAEQEPSCWGHPLSRAATQNTQSMPSQSLSSVQDYGKRLKANLKNNTQT 120
Qy 120 GPALGRRPPWPLGRASKASTPKPGTGPVPSFAEKVSDPEPPQPPRPGRLQHLQASL 179
Db 121 GPTQSRKLIQKRSLSLTVPAPEPPGSKTESPCDEADDDALPRVPEPRPLGQQLKRSLL 180
Qy 180 SQRLGSLDQGLWLRCHSEVDPFLGAPKACRPDLGSEESQLLIPGESAVILPGAGS----Q 235
Db 181 SRRLTSLDQGLWLRCHNRVSDLLLEVGCGLDLASAEQSPQMSGKVNIADPIQSEVSQ 240
Qy 236 GPASAFQEVSVIRVSGPQPSGSGEKRRWNEPWPESPAQVQOESSQAGPPSEGAGAVAE 295
Db 241 SPEATIAQQAQVLSQSPKSIKSKGRKWKNEK-GEDFAQ-DQPSGAGPLSEGARATVHG 298
```

```

Qy 296 EDPGCEPVOAOPPCSSPSNRYHGLSPSSQARAGKAGTAPLHIPPRLARHGRNVR 355
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 QDPFCEPTQVNVQPCNS-----SNQARTKAKGTHLHSPASPLDRGNVIR 347
Qy 356 LNMOKHYVRGRALRSLLRQAWKQKWKKEGCGGGAATVTKESFLNEQDHWAAQ 415
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
348 LNMKNKRVVRGANRGLLRKQVWKQKWKQAFGGSGPRATDKDTCFCGQGFCHWASQ 407
Qy 416 CPRPA-----SEEDTDAVGP-----EPLVPSQPVPPEVPSL 446
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
408 CSQPGPTLTQVBEGRDDKQPTSTLEEVAQRTGTASCHSGEETQPAAPELQ-VPHCPT- 465
Qy 447 DPTVLPYLSLSPGQLAETPAEVAOLQOLHQAQPRQOERAVRMLISGISTLVLPTGA 506
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
466 --PMSPLYPGPGQVETPAEVAOLQOLHQAQPRQOERAVRMLISGISTLVLPTGA 523
Qy 507 GKSCLYQPLALLYSRRSCLTLVSPILSLMDQVSGLPCLKAACIHSGMTKQRESVL 566
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
524 GKSCLYQPLALLYAQRSCLTLVSPILSLMDQVSDLPCLKAACIHSGMTKQRESVL 583
Qy 567 QKIRAAQVHVLMLTPEALVGA-----GGLPPAAQLPPVAFACIDBAHCLSMNSHNFRCY 621
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
584 KKVRAAQVHVLIVSPEALVGCARGPSLPOAAQLPPIAFACIDBHVHCLSQWSHNFRCY 643
Qy 622 LRVCKVLRERMGVHCFGLGTATARTASDVAAHLAVAEEDPLHGPAPVPTNLHLSVMD 681
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
644 LRVCKVLRERMGVHCFGLGTATARTASDVAAHLAVAEEDPLHGPAPVPTNLHLSVMD 703
Qy 682 RDTQALLTLQGRFONLDSIIICNRRDTERIAALLRTCLHAAWVPGSGRPAKTTA 741
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
704 RDSQALLVTLLQGRFRLDSVIICTHRTQNGWALLRLCLSMVDSRPGCGPEAIA 763
Qy 742 EAYHAGMCSRERRRVRAPMOCQLRVVATVAFMGMLDRPDVRAVLHLGLPPSFESYQA 801
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
764 EAYHAGMSSQERRRVQAFMRGHLRMVATVAFMGMLDRPDVRAVLHLGLPPSFESYQA 823
Qy 802 VBRAGDQOPACHLFLQOGEDLRELRRHVHADSTDPLAVKRLVORVPPACTCTCTRP 861
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
824 IGRAGRDGKPAHCHLFMHPQGEDLWELRRHAHADSTDPLAVKRLVORVPPACTCS----- 878
Qy 862 SBOGAVGERPVKVPPOEABOLS-HQAAPGRVRCVCHGERALPIQLTVQALDMPPEAI 920
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
879 -----QRPVKSSEPEVEKHSQOQTYFVLGOACLCHGERALPVQSTVQALDMPPEAI 929
Qy 921 ETLCLYLELPHHLELLATTTTHCLNCPGPAQLOALAHRCPPELAVCLAQOLPEDPCQ 980
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
930 ETLCLYLELPHHLELLPWTYAQCCHLCLGGSQAQLOALAHRCPPELAAQAKWPKDTSQ 989
Qy 981 GSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRGTGVLVEFSELAFLHRS 1040
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
990 GRSSLEFGVVELADSMGWKLASVROALHOLKWDPEPKGAAQGTGVLVKFSELAFLHRS 1049
Qy 1041 GDLTAREKDOI CDLYGRVQARERQALRLRTFOAHSVAFPSGPCLEQDDEBSRSL 1100
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1050 GDLTAREKDOI CDLYGRVQARERQALRLRTFOAHSVAFPSGPCLEQDDEBSRSL 1109
Qy 1101 KDLGRYFEEBEGQPGMEDAQGPQOARLQDWEQVRCDIROFLSLRPEKFSRRAV 1160
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1110 KTLVSYFEEB-EREETWTDQGPQGTQLQDWEQVRCDIROFLSLRPEKFSRRAV 1168
Qy 1161 ARIFHIGSPCYPAQVYGDORFWRKYLHLSFHALVGLATELL 1204
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1169 ARIFHIGSPCYPAQVYGLDRFRWRKYLHLSFHALVGLATELL 1212

```

RESULT 6

US-11-097-143-16743
; Sequence 16743, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

```

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16743
; LENGTH: 1579
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-16743

```

```

Query Match      26.8%; Score 1723.5; DB 6; Length 1579;
Best Local Similarity 28.9%; Pred. No. 2.1e-103;
Matches 467; Conservative 217; Mismatches 459; Indels 471; Gaps 45;

Qy 10  RLQAWERAFRRQRGRPSQDDVEAAPBETRALYREYTLKKT----- 51
Db 16  RVKWEKDFKKKGRVPSKYDIRDASQEIIRDSYKMYKLTSTFLEETLNDVLSDEGYDIL 75

Qy 52  -TQAGG-----GLRS-----ESLP----- 66
Db 76  EMSQASDFGVSMLDQDVSLNEGQPLDLISALVQSSGNLEEPQSVESGFSNLDLDPN 135

Qy 67  -----AAAEAP-EPRCWGPHLNAATKSPQ-----TPGRSRQ 100
Db 136  ROVLTNLVNRDENHVIRKFEAVEBELPINQAWGLNVSKPPAPPQVPEASKAFGHQKQ- 194

Qy 101  SVDPYQRLKANLKTGL-QAGPALGRR--PWPLGRASSKASTPKPGTGPVPSFAKVSVD 157
Db 195  --PRAGASLKPSSLAKLFPQSSRGFAKRNPRKPLSRCVSSSTTSLSSVPTDHHBELDF 252

Qy 158  EPPQLPEQPPRGRILQ-----HLQASLQRLGSLDPGNLOR-----CHSEVP- 199
Db 253  ETILIRKAQYKKEKQQAATANNPMLASEHSKSIKT---LVDDGMLRRNTKENTLDEEVPF 309

Qy 200  ---DFLGAPK-----ACRPDLGSEESQLLIPGESAVL-----GP 230
Db 310  ABANNSTGSKTTFNGLANLDLSKLPTVKEKVLQAKPDQWAIQELQTDNWSNMNQK 369

Qy 231  GAGSQGPASAFQEYISIRVSGPQSS-----SGGEKRRWNEE----- 267
Db 370  DHLNHTTPASSQKSVAPK-NKPPPEQETDSDSDVSAESEEQEPQVYRQLSKRKIVS 428

Qy 268  -----PWESPAQVOQES---SQAGP----- 284
Db 429  TASGKVEVAAPVEIPNKPVEPETETFAQENPDFSADEQDATVYPENKKDKAKRQAAAG 488

Qy 285  -----PSEGAGAVAVEEDPPGPVQAOQPPQ--- 310
Db 489  QKTKPKAEPKTEKKAQVKAEPKPKPRNSKKAIAVEPAPDPPEDEKQPLNPEDL 548

Qy 311  -----CSSPSNRYH-----GLSPSSQARAGKAGTAP 338
Db 549  KYVLALAGDITSVPRINVQDLEADATAQRYIRTFRTAAGPNPGLSEGSNIRVDEKRAAR 608

```



```
Db 498 FVDMRSIIYCKFQAEITDFVSKYL---CDN-----NITAKSVHSGLLKNRSRV 543
Qy 757 QRAFMQGLRVVAVTAVFGMGLDRPDVRAVLHLGLPPSFESYQAVGRAGRGOPAHCHL 816
Db 544 QELFCSNKIRVP-----SY-----NF 559
Qy 817 FLOPQEDLRLRHVHADSTDFLAVKLVQVFPACTCTTRPPSEOGAVGGERVPVK 876
Db 560 TEQP-----FLGSDGVDGYAMSKFLYQIF-----SSENTTG-----590
Qy 877 YPPQAEQLHQADPGRRCVCMGHERALPIQLTVQALDMPBEAIEITLLCYLELHPHWLE 936
Db 591 -----CIG-----SLAKELTSKFDEKEVLTILTQLEIGDQOYIR 627
Qy 937 LIATTYTTCRLNCPGPAQLQALAHRCPPPLAVC-----LAQQLPEDPGQSSSEVFDVMKL 992
Db 628 LLPQPSVTCVL-----YFKTSPQLLADKDLILRSVLRNRMKDGHYVFDIPRI 676
Qy 993 VDSMGWELASVRRALCQOWDHEPTGVRGTGVLVERS-ELAFHLRSPG-----1041
Db 677 ANDLKITMNEV-----FDHLK-----LKFSGEISFELKDPAYCYVILWRPD 718
Qy 1042 DUTAEKQICDLYGRVQARERQALARRTFQAFHSVAPPSC-----GPCLE 1090
Db 719 DFNALSAN-----LTKWISEVSSKISKLDAMP-ALANPAVKGCKRTGCGSSQHTPCIQ 772
Qy 1091 QODBERSTRKDLGLGRYFEEERGQPPGMDAQQPEPQARLDWEDQVRCDIROFLSLR 1150
Db 773 KKIME-----YFSKDDGTSEND-----RTLQKSPFLQADIKVFIQSN 812
Qy 1151 PREKSSRAVARIFIGISGPCYPAQVYQDPRFWRKYLHLSFHALVGLATELLQV 1206
Db 813 SPAKTPRAVARIMEGISSPAPFSPVTSKN-HFWGRYVEVDFFLYMAKAELVKL 867

RESULT 8
US-10-374-077-75
; Sequence 75, Application US/10374077
; Publication No. US20040006779A1
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Oshima, Junko
; Mulligan, John T.
; Schellenberg, Gerald D.
; TITLE OF INVENTION: ANTIBODIES AGAINST GENE PRODUCTS RELATED TO
; WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,077
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Stephen
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 100107.401D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 607 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-10-374-077-75

Query Match Similarity 9.8%; Score 632.5; DB 4; Length 607;
Best Local Similarity 37.6%; Pred. No. 1.2e-32; Indels 75; Gaps 15;
Matches 164; Conservative 63; Mismatches 134;

Qy 474 BOLGQAFRPOQERAVNRVIRIISGISTLVLPFGAGKSLCYQLPALIYSRRSPCLTIIVSPL 533
Db 18 ETFGYQFRFQOEIIDTVLSGRDCLVVMPTGGKSLCYQIIPALLNG----LTVVVSPL 73
Qy 534 LSLMDDQVSGLPP-CLKAACTHSGMTRKQRESVLQKRAAQVHVLMLTPEALVGAGGLPP 592
Db 74 ISLMKDQVQLQANGVAAACLNSTQTRQQLLEVMTGCRGQIRLLYIAPERLMDNFLEH 133
Qy 593 AAQLPPVAFACIDEAHCLSQSHNFRPCYLKVCVLRBMGVHCFGLGTATATRTASDV 652
Db 134 LAHNPVLLA-VDEAHCTISQWGHDFRPEYALGQ-LRQRFPTLPPMALATATADDTTODI 191
Qy 653 AQLHVAEEDPLHGPAPVPTNLHLSVSMDDRTDQALLTLQCKRFQNLDSI-----703
Db 192 VRLG-----LNDPL-----IQIS-SFDRNIRYMLM-----EKEKPLDQLMRYVQEOG 235
Qy 704 ---IYCNRRDTERIALALRTCLHAAWVPGSGRPAKTTABAYHAGMCSRERRRVQRAF 760
Db 236 KSGIYCNRAKVEDTAAL-----QSGISAAAVHAGLENNVRADVQEK 281
Qy 761 MQQLRVVAVTAVFGMGLDRPDVRAVLHLGLPPSFESYQAVGRAGRGOPAHCHLFL-- 818
Db 282 QRDDLQIVVATVAFGMGINKENRVFVHFDPNIESYQYGTGRAGRDGLPAEAMLFYDP 341
Qy 819 -----OPQGEDLRELRRH---VHADSTDTFLAVKRLV-----QRVFPACTC- 855
Db 342 ADMAWLRRCLEKPOGQ-LQDIERHKLNAMGAFABQAQTCRLRLVLLNYFEGEQEPCGNC 400
Qy 856 TCTRPPSEQEGAVGGE 871
Db 401 ICLDPPKQYDGSDDAQ 416

RESULT 9
US-09-815-242-10414
; Sequence 10414, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; PROKARYOTES
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
```

```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10414
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10414

Query Match
Best Local Similarity 9.8%; Score 631.5; DB 3; Length 610;
Matches 164; Conservative 63; Mismatches 134; Indels 75; Gaps 15;

Qy 474 EQLGHQAFRPGQBRVAVNRILSGISTLVLPTGAGKSLCYQLPALLYRRSPCLTLVVSPL 533
Db 21 ETFGYQFRPGQBEIITVLGRDCLVMPGTGGKSLCYQIPALLNG----LTVVVSPL 76

Qy 534 LSLMDDQVSGLPP-CLKAACIHSGMTRKQRESVLQKIRAAQVHVLMLTPEALVAGAGLPP 592
Db 77 ISLMKDQVDQLQANGVAAACLNSTQREQLLEVMTGCTGQIRLLYIAPERLMDNLFLEH 136

Qy 593 AAOPLPVAFACIDEAHCLSQSHNFRPCYLKRVCKLVRMGVHCFLGLTATATRRRTASDV 652
Db 137 LAHNVPVLLA-VDEAHCSIQWGHDFRPEYALGQ-LRQRFPTLPFMALTATADTTTQDI 194

Qy 653 AQHLAVAEEDPLHGPAPVPTNLHLSVSMRDRTDQALLTLQGRFQNLDSI----- 703
Db 195 VRLIG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKFKPLDQLMRVYQSQRG 238

Qy 704 ---IYCNRRDTERIAALLTCLHAAWVPGSGRAPKTTABAYHAGMCSRERRVQRAF 760
Db 239 KSGIYCNRAKVEDTAARL-----QSKGISAAAVHAGLENNVRADVQEKF 284

Qy 761 MQQLRVVATVAFGMGLDRPDVRAVLHLGLPPSFESYVQAVGRAGRDGPAHCHLFL-- 818
Db 285 QRDDLQIVVATVAFGMGINKPNRVFVHFDPINRISYYOETGRAGRDGLPAEAMLFYDP 344

Qy 819 -----QFGEDLRELRRH---VHADSTDFLAVKRLV-----QRVFPACTC- 855
Db 345 ADMAWLRCLLEKEPQGO-LQDIERHKLNAMGAFEAOTCRLVLLNYFGEGRQPCGNC 403

Qy 856 TCTRPPEQEGAVGGE 871
Db 404 ICLDPPKQYDGSSTDAQ 419
```

```
RESULT 10
US-10-282-122A-43357
; Sequence 43357, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zveskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsythe, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43357
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43357

Query Match
Best Local Similarity 9.8%; Score 631.5; DB 4; Length 610;
Matches 164; Conservative 63; Mismatches 134; Indels 75; Gaps 15;

Qy 474 EQLGHQAFRPGQBRVAVNRILSGISTLVLPTGAGKSLCYQLPALLYRRSPCLTLVVSPL 533
Db 21 ETFGYQFRPGQBEIITVLGRDCLVMPGTGGKSLCYQIPALLNG----LTVVVSPL 76

Qy 534 LSLMDDQVSGLPP-CLKAACIHSGMTRKQRESVLQKIRAAQVHVLMLTPEALVAGAGLPP 592
Db 77 ISLMKDQVDQLQANGVAAACLNSTQREQLLEVMTGCTGQIRLLYIAPERLMDNLFLEH 136

Qy 593 AAOPLPVAFACIDEAHCLSQSHNFRPCYLKRVCKLVRMGVHCFLGLTATATRRRTASDV 652
Db 137 LAHNVPVLLA-VDEAHCSIQWGHDFRPEYALGQ-LRQRFPTLPFMALTATADTTTQDI 194

Qy 653 AQHLAVAEEDPLHGPAPVPTNLHLSVSMRDRTDQALLTLQGRFQNLDSI----- 703
Db 195 VRLIG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKFKPLDQLMRVYQSQRG 238

Qy 704 ---IYCNRRDTERIAALLTCLHAAWVPGSGRAPKTTABAYHAGMCSRERRVQRAF 760
Db 239 KSGIYCNRAKVEDTAARL-----QSKGISAAAVHAGLENNVRADVQEKF 284

Qy 761 MQQLRVVATVAFGMGLDRPDVRAVLHLGLPPSFESYVQAVGRAGRDGPAHCHLFL-- 818
Db 285 QRDDLQIVVATVAFGMGINKPNRVFVHFDPINRISYYOETGRAGRDGLPAEAMLFYDP 344

Qy 819 -----QFGEDLRELRRH---VHADSTDFLAVKRLV-----QRVFPACTC- 855
Db 345 ADMAWLRCLLEKEPQGO-LQDIERHKLNAMGAFEAOTCRLVLLNYFGEGRQPCGNC 403

Qy 856 TCTRPPEQEGAVGGE 871
Db 404 ICLDPPKQYDGSSTDAQ 419
```

```
RESULT 11
US-09-753-143-77
; Sequence 77, Application US/09753143
; Patent No. US20020102550A1
; GENERAL INFORMATION:
; APPLICANT: NATHAN A. ELLIS, JAMES GERMAN, AND JOANNA
; GRODEN
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
```


Search completed: December 27, 2005, 22:12:34
Job time : 184.4 secs

RESULT 15

US-10-282-122A-74973

Sequence 74973, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELI/PRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:38:41 ; Search time 54.2 Seconds
(without alignments)
2448.199 Million cell updates/sec

Title: US-09-889-325-4_COPY_907_1208

Perfect score: 1617
Sequence: 1 QLTVALDMPREAIETLLCY.....HLSFHALVGLATELLOVAR 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 1617 | 100.0 | 1142 | 8 | Abm84821 Human dia |
| 2 | 1617 | 100.0 | 1144 | 7 | Add01202 Human nuc |
| 3 | 1617 | 100.0 | 1208 | 2 | Aaw95050 Human hel |
| 4 | 1617 | 100.0 | 1208 | 3 | Aab20993 Human Rec |
| 5 | 1617 | 100.0 | 1208 | 5 | Abg933387 Human Rec |
| 6 | 1617 | 100.0 | 1208 | 8 | Adq21562 Human sof |
| 7 | 1617 | 100.0 | 1208 | 8 | Adu80701 Human RTS |
| 8 | 1594.5 | 98.6 | 1306 | 8 | Adm90956 Human pha |
| 9 | 1144.5 | 70.8 | 1216 | 9 | Adw44021 Mouse REC |
| 10 | 427 | 26.4 | 1579 | 4 | ABBG63317 |
| 11 | 145.5 | 9.0 | 206 | 8 | ADY13119 Plant ful |
| 12 | 108.5 | 6.7 | 762 | 7 | ABO80677 Pseudomon |
| 13 | 107 | 6.6 | 751 | 6 | ABR82253 Human nuc |
| 14 | 107 | 6.6 | 755 | 8 | ABM80914 Tumour-as |
| 15 | 107 | 6.6 | 1029 | 3 | Aay96736 PRO3434, |
| 16 | 107 | 6.6 | 1029 | 3 | Aab24027 Human PRO |
| 17 | 107 | 6.6 | 1029 | 5 | ABG927710 Human sec |
| 18 | 107 | 6.6 | 1029 | 5 | Aau77765 Tumour as |
| 19 | 107 | 6.6 | 1029 | 5 | ABG91362 Novel hum |
| 20 | 107 | 6.6 | 1029 | 5 | ABG31402 Human PRO |
| 21 | 107 | 6.6 | 1029 | 6 | ABU72378 Novel hum |
| 22 | 107 | 6.6 | 1029 | 6 | ABU80868 Human sec |
| 23 | 107 | 6.6 | 1029 | 6 | ABG73315 Human PRO |
| 24 | 107 | 6.6 | 1029 | 6 | ABU60814 Human sec |

| | | | | | |
|----|-------|-----|------|---|--------------------|
| 25 | 107 | 6.6 | 1029 | 6 | ABU81237 Human PRO |
| 26 | 107 | 6.6 | 1029 | 6 | ABU82958 Human PRO |
| 27 | 107 | 6.6 | 1029 | 7 | ABO01895 Novel hum |
| 28 | 107 | 6.6 | 1029 | 7 | ADB89026 Human tum |
| 29 | 107 | 6.6 | 1029 | 7 | AAE38827 Human PRO |
| 30 | 107 | 6.6 | 1029 | 7 | ADJ58520 Human PRO |
| 31 | 107 | 6.6 | 1029 | 8 | ADI10343 Human PRO |
| 32 | 107 | 6.6 | 1029 | 9 | AEb47258 Human tum |
| 33 | 107 | 6.6 | 1029 | 9 | AEb70219 Human tum |
| 34 | 107 | 6.6 | 1377 | 6 | ABP56518 Human BAA |
| 35 | 107 | 6.6 | 2190 | 7 | ADB80326 Human MDD |
| 36 | 106 | 6.6 | 253 | 8 | ADQ66277 Novel hum |
| 37 | 105.5 | 6.5 | 699 | 7 | ADI21211 Novel hum |
| 38 | 102.5 | 6.3 | 279 | 7 | ABO70187 Pseudomon |
| 39 | 102 | 6.3 | 817 | 7 | ABO77141 Pseudomon |
| 40 | 100 | 6.2 | 552 | 4 | AAE01117 Human gen |
| 41 | 100 | 6.2 | 552 | 5 | ABG64586 Human alb |
| 42 | 100 | 6.2 | 552 | 8 | ADL77853 Albumin f |
| 43 | 99.5 | 6.2 | 512 | 7 | ABO73593 Pseudomon |
| 44 | 98 | 6.1 | 719 | 4 | AAM00798 Human bon |
| 45 | 98 | 6.1 | 724 | 4 | AAM00911 Human bon |

ALIGNMENTS

RESULT 1
ABM84821
ID ABM84821 standard; protein; 1142 AA.
XX
AC ABM84821;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:5070.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patory S, Shi X, Suarez CJ;
XX
DR WPI: 2004-329368/30.
DR N-PSDB; ACN43473.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human

CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX

SQ Sequence 1142 AA;

Query Match 100.0%; Score 1617; DB 8; Length 1142;
Best Local Similarity 100.0%; Pred. No. 1.1e-158;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QLTVQALDMPPEAEITLLCYLELPHHLELLATTYTHCRNLCPGPAQQAALAHRCPPPL 60
Db 841 QLTVQALDMPPEAEITLLCYLELPHHLELLATTYTHCRNLCPGPAQQAALAHRCPPPL 900
Qy 61 AVCLAQQLPEDPGQSSSVEFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRGTGV 120
Db 901 AVCLAQQLPEDPGQSSSVEFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRGTGV 960
Qy 121 LVFESELAHLRSPGDLTAEKDKQICDPLYGRVQARERQALRLRRTFOAFHSAVPSCG 180
Db 961 LVFESELAHLRSPGDLTAEKDKQICDPLYGRVQARERQALRLRRTFOAFHSAVPSCG 1020
Qy 181 PCLEQQDEERSTRKDLLGRYFEEEGQEPGMEADQGPQARLQDWDQVRCDIROF 240
Db 1021 PCLEQQDEERSTRKDLLGRYFEEEGQEPGMEADQGPQARLQDWDQVRCDIROF 1080
Qy 241 LSLRPEKFSRAVARIFHIGISPCYPAQVYGQDRFRWRKYLHLSFHALVGLATEELLQV 300
Db 1081 LSLRPEKFSRAVARIFHIGISPCYPAQVYGQDRFRWRKYLHLSFHALVGLATEELLQV 1140
Qy 301 AR 302
Db 1141 AR 1144

RESULT 2
ADD01202
ID ADD01202 standard; protein; 1144 AA.
XX
AC ADD01202;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human nucleic acid-associated protein NAAP-40 SEQ ID NO:40.
XX
KW human; nucleic acid-associated protein; NAAP; cytostatic;
KW antiarteriosclerotic; anti-HIV; antiallergic; cerebroprotective;
KW antiparkinsonian; anticonvulsant; neurotropic; neuroprotective;
KW antiinflammatory; ophthalmological; thyromimetic; antiarthritic;
KW hepatocytic; antibacterial; virucide; protozoacide; antiparasitic;
KW fungicide; gene therapy; cell proliferative disease; cancer;
KW atherosclerosis; hepatitis; neurological disorder; Parkinson's disease;
KW Alzheimer's disease; stroke; epilepsy; developmental disorder;
KW renal tubular acidosis; anaemia; glaucoma; hypothyroidism;
KW autoimmune disorder; inflammatory disorder; AIDS; allergy;
KW atopic dermatitis; arthritis; infection.
XX
OS Homo sapiens.
XX
FN WO2003054219-A2.
XX
PD 03-JUL-2003.
XX

PF 18-DEC-2002; 2002WO-US041115.
XX
PR 19-DEC-2001; 2001US-0343004P.
PR 11-JAN-2002; 2002US-0347633P.
PR 25-JAN-2002; 2002US-0351749P.
PR 22-FEB-2002; 2002US-0359498P.
XX
(INCY-) INCYTE GENOMICS INC.
XX
PI Azimzi Y, Baughn MR, Becha SD, Borowsky ML, Burford N;
PI Elliott VS, Emerling BM, Forsythe IO, Gorvad AE, Griffin JA;
PI Kable AE, Khare R, Lal PG, Lee EA, Lee SY, Li JX, Marquis JP;
PI Ramkumar J, Richardson TW, Sprague WW, Swarnakar A, Tang YT;
PI Chawla NK, Warren BA, Yue H;
XX
DR WPI; 2003-559157/52.
DR N-PSDB; ADD01259.
XX
PT New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
PT infections.
XX
PS Claim 1; SEQ ID NO 40; 405pp; English.
XX
CC The present invention describes human nucleic acid-associated proteins
CC designated NAAP-1 to NAAP-57. The human NAAPs have cytostatic,
CC antiarteriosclerotic, anti-HIV, antiallergic, cerebroprotective,
CC antiparkinsonian, anticonvulsant, neurotropic, neuroprotective,
CC antiinflammatory, ophthalmological, thyromimetic, antiarthritic,
CC hepatocytic, antibacterial, virucide, protozoacide, antiparasitic and
CC fungicide activities, and can be used in gene therapy. The NAAP protein
CC and polynucleotide sequences can be used in diagnosing, treating and
CC preventing diseases or conditions associated with the decreased
CC expression or overexpression of NAAP, such as cell proliferative diseases
CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
CC hypothyroidism), autoimmune/inflammatory disorders (AIDS, allergies,
CC atopic dermatitis, arthritis) and infections (e.g. bacterial, viral,
CC parasitic, protozoal, fungal). The present sequence represents human NAAP
CC -40, from the present invention.
XX
SQ Sequence 1144 AA;

Query Match 100.0%; Score 1617; DB 7; Length 1144;
Best Local Similarity 100.0%; Pred. No. 1.1e-158;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QLTVQALDMPPEAEITLLCYLELPHHLELLATTYTHCRNLCPGPAQQAALAHRCPPPL 60
Db 843 QLTVQALDMPPEAEITLLCYLELPHHLELLATTYTHCRNLCPGPAQQAALAHRCPPPL 902
Qy 61 AVCLAQQLPEDPGQSSSVEFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRGTGV 120
Db 903 AVCLAQQLPEDPGQSSSVEFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRGTGV 962
Qy 121 LVFESELAHLRSPGDLTAEKDKQICDPLYGRVQARERQALRLRRTFOAFHSAVPSCG 180
Db 963 LVFESELAHLRSPGDLTAEKDKQICDPLYGRVQARERQALRLRRTFOAFHSAVPSCG 1022
Qy 181 PCLEQQDEERSTRKDLLGRYFEEEGQEPGMEADQGPQARLQDWDQVRCDIROF 240
Db 1023 PCLEQQDEERSTRKDLLGRYFEEEGQEPGMEADQGPQARLQDWDQVRCDIROF 1082
Qy 241 LSLRPEKFSRAVARIFHIGISPCYPAQVYGQDRFRWRKYLHLSFHALVGLATEELLQV 300
Db 1083 LSLRPEKFSRAVARIFHIGISPCYPAQVYGQDRFRWRKYLHLSFHALVGLATEELLQV 1142
Qy 301 AR 302
Db 1143 AR 1144

RESULT 3
AAW95050
ID AAW95050 standard; protein; 1208 AA.

XX AC AAW95050;
XX DT 14-MAY-1999 (first entry)
XX DE Human helicase protein.

XX KW RecQ4 gene; helicase; Werner's syndrome; Bloom's syndrome; human.

XX OS Homo sapiens.
XX PN WO9905284-A1.
XX PD 04-FEB-1999.

XX PF 10-JUL-1998; 98WO-JP003114.

XX PR 25-JUL-1997; 97JP-00200387.

XX PA (AGEN-) AGENE RES INST CO LTD.

XX PI Shimamoto A, Kitao S, Furuichi Y;

XX DR WPI; 1999-142939/12.

XX DR N-PSDB; AAX21656.

XX PT New human helicase gene RecQ4 - used for investigation and diagnosis of

XX PT helicase-implicated diseases such as Werner's syndrome.

XX PS Claim 1; Page 35-42; 67pp; Japanese.

XX CC The present sequence represents a protein having helicase activity
XX CC encoded by the human gene RecQ4. The gene has significant homology to the
XX CC Escherichia coli helicase gene (RecQ). Host cells transformed with
XX CC vectors comprising the RecQ4 gene are used for the recombinant expression
XX CC of the protein. The gene may be used for the study and diagnosis of
XX CC disorders in which helicase activity is involved, such as Werner's and
XX CC Bloom's syndromes in which mutations in the helicase gene are implicated

XX SQ Sequence 1208 AA;

Query Match 100.0%; Score 1617; DB 2; Length 1208;
Best Local Similarity 100.0%; Pred. No. 1.2e-158;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLTVALDMPPEAIEITLLCYLELHPHMLELLATTYTHCRINCPCGGPAQLQALAHRCPPPL 60
DB 907 QLTVALDMPPEAIEITLLCYLELHPHMLELLATTYTHCRINCPCGGPAQLQALAHRCPPPL 966

QY 61 AVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRGTGTV 120
DB 967 AVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRGTGTV 1026

QY 121 LVFSELAFLHRS PGDLTAEKDKQICDFLYGRVQARERQALRLRTFOAFHSVAFPPSCG 180
DB 1027 LVFSELAFLHRS PGDLTAEKDKQICDFLYGRVQARERQALRLRTFOAFHSVAFPPSCG 1086

QY 181 PCLEQDDEERSTRLKDLGRYFEFEEBGPQGMEDAQGPQCARLQDWDQVRCDIQF 240
DB 1087 PCLEQDDEERSTRLKDLGRYFEFEEBGPQGMEDAQGPQCARLQDWDQVRCDIQF 1146

QY 241 LSLRPEEKSSAVARIIFGIGSPCYPAQVYQDQRRFRWKYHLHSHALVGLATELLQV 300
DB 1147 LSLRPEEKSSAVARIIFGIGSPCYPAQVYQDQRRFRWKYHLHSHALVGLATELLQV 1206

QY 301 AR 302

DB 1207 AR 1208

RESULT 4

AAW20993
ID AAW20993 standard; protein; 1208 AA.

XX AC AAW20993;

XX DT 11-DEC-2000 (first entry)

XX DE Human RecQ4 helicase.

XX KW RecQ4 helicase; human; Rothmund-Thomson syndrome; chromosome 8q24.3;
XX KW poikiloderma congenitale; autosomal recessive; skin disorder;
XX KW dermatology; antibody; prenatal diagnosis; gene therapy.

XX OS Homo sapiens.

XX PN WO200043522-A1.

XX PD 27-JUL-2000.

XX PF 19-JAN-2000; 2000WO-JP000233.

XX PR 19-JAN-1999; 99JP-00011218.

XX PA (AGEN-) AGENE RES INST CO LTD.

XX PI Kitao S, Shimamoto A, Furuichi Y;

XX DR WPI; 2000-524241/47.

XX DR N-PSDB; AAW72320, AAW72321.

XX PT RecQ4 helicase gene, gene products and antibody, used in the diagnosis
XX PT and treatment of Rothmund-Thomson syndrome, e.g. by gene therapy.

XX PS Example 5; Page 83-92; 115pp; Japanese.

XX CC The present sequence represents human RecQ4 helicase. The invention
XX CC relates to the genomic DNA sequence of human RecQ4 helicase (AAW72320).
XX CC Mutations in the RecQ4 helicase gene, located on chromosome 8q24.3, are
XX CC the cause of Rothmund-Thomson syndrome (also known as poikiloderma
XX CC congenitale), an autosomal recessive skin disorder principally occurring
XX CC in females and often accompanied by juvenile cataracts, saddle nose,
XX CC congenital bone defects, hypogonadism and disturbances in the growth of
XX CC hair, nails and teeth. The invention also relates to vectors and host
XX CC cells comprising the human RecQ4 helicase genomic sequence. It
XX CC additionally encompasses use of the RecQ4 helicase protein as a
XX CC therapeutic and anti-RecQ4 antibodies as diagnostic agents. The RecQ4
XX CC helicase gene and its products, and anti-RecQ4 helicase antibodies are
XX CC useful in the diagnosis, especially prenatal diagnosis, and treatment of
XX CC Rothmund-Thomson syndrome. The genomic sequence may especially be used in
XX CC gene therapy for this condition

XX SQ Sequence 1208 AA;

Query Match 100.0%; Score 1617; DB 3; Length 1208;
Best Local Similarity 100.0%; Pred. No. 1.2e-158;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLTVALDMPPEAIEITLLCYLELHPHMLELLATTYTHCRINCPCGGPAQLQALAHRCPPPL 60
DB 907 QLTVALDMPPEAIEITLLCYLELHPHMLELLATTYTHCRINCPCGGPAQLQALAHRCPPPL 966

QY 61 AVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRGTGTV 120
DB 967 AVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRGTGTV 1026

QY 121 LVFSELAFLHRS PGDLTAEKDKQICDFLYGRVQARERQALRLRTFOAFHSVAFPPSCG 180
DB 1027 LVFSELAFLHRS PGDLTAEKDKQICDFLYGRVQARERQALRLRTFOAFHSVAFPPSCG 1086

QY 181 PCLEQDDEERSTRLKDLGRYFEFEEBGPQGMEDAQGPQCARLQDWDQVRCDIQF 240
DB 1147 PCLEQDDEERSTRLKDLGRYFEFEEBGPQGMEDAQGPQCARLQDWDQVRCDIQF 1206

Db 1087 PCLEQDEERSRTRKDLLGRYFEEEGQPGMEDAQGPQARLQDWDQVRCDIRQF 1146
Qy 241 LSLRPEKFSRAVARIFHIGIGSPCYPAQVYGQDRFRWRYLHLSFHALVGLATEELLQV 300
Db 1147 LSLRPEKFSRAVARIFHIGIGSPCYPAQVYGQDRFRWRYLHLSFHALVGLATEELLQV 1206
Qy 301 AR 302
Db 1207 AR 1208

RESULT 5
ABG93387
ID ABG93387 standard; protein; 1208 AA.
XX
AC ABG93387;
XX
DT 20-NOV-2002 (first entry)
XX
DE Human RecQ protein-like 4 (RECQL4).
XX
KW Human; RecQ protein-like 4; RECQL4; infection; inflammation; cytostatic;
KW tumour formation; cancer; antiinflammatory; antimicrobial;
KW antisense therapy.
XX
OS Homo sapiens.
XX
FN US6436706-B1.
XX
PD 20-AUG-2002.
XX
PF 23-FEB-2001; 2001US-00792594.
XX
PR 23-FEB-2001; 2001US-00792594.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ward DT, Watt AT;
XX
WPI; 2002-689941/74.
XX
N-PSDB; ABS68851.
XX

New antisense compounds targeted to nucleic acids encoding RecQ protein-like 4, useful for modulating expression of the nucleic acid and treating diseases associated with expression of the nucleic acid in humans.
Disclosure; Col 47-56; 45pp; English.

The invention relates to a compound targeted to specific nucleobases of RecQ protein-like 4 (RECQL4) and which hybridises and inhibits the expression of RECQL4. The compound is useful for inhibiting the expression of RECQL4 in cells or tissues and for treating an animal, particularly a human suspected of having or being prone to a disease or condition associated with expression of RECQL4. The compound is useful for diagnostics, therapeutics and as a research reagent, e.g. prophylactically to prevent or delay infection, inflammation or tumour formation. This sequence represents human RECQL4 polypeptide

Sequence 1208 AA;
Query Match 100.0%; Score 1617; DB 5; Length 1208;
Best Local Similarity 100.0%; Pred. No. 1.2e-158;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLTVOALDMPPEAEITLLCYLELHPHHLATYTHCRINCPCGGPAQQAQALAHRCPL 60
Db 907 QLTVOALDMPPEAEITLLCYLELHPHHLATYTHCRINCPCGGPAQQAQALAHRCPL 966
Qy 61 AVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRRTGV 120
Db 967 AVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRRTGV 1026
Qy 121 LVEFSELAFLHRSFGDLTAEKQICDFLYGRVQARERQALRLRRTQAFHSAVPSCG 180

Db 1027 LVEFSELAFLHRSFGDLTAEKQICDFLYGRVQARERQALRLRRTQAFHSAVPSCG 1086
Qy 181 PCLEQDEERSRTRKDLLGRYFEEEGQPGMEDAQGPQARLQDWDQVRCDIRQF 240
Db 1087 PCLEQDEERSRTRKDLLGRYFEEEGQPGMEDAQGPQARLQDWDQVRCDIRQF 1146
Qy 241 LSLRPEKFSRAVARIFHIGIGSPCYPAQVYGQDRFRWRYLHLSFHALVGLATEELLQV 300
Db 1147 LSLRPEKFSRAVARIFHIGIGSPCYPAQVYGQDRFRWRYLHLSFHALVGLATEELLQV 1206
Qy 301 AR 302
Db 1207 AR 1208

RESULT 6
ADQ21562
ID ADQ21562 standard; protein; 1208 AA.
XX
AC ADQ21562;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4382.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
WPI; 2004-441208/41.
XX

Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

Example 2; SEQ ID NO 4382; 210pp; English.

The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated protein of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

Sequence 1208 AA;
Query Match 100.0%; Score 1617; DB 8; Length 1208;
Best Local Similarity 100.0%; Pred. No. 1.2e-158;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLTVOALDMPPEAEITLLCYLELHPHHLATYTHCRINCPCGGPAQQAQALAHRCPL 60

Db 907 QLTVALDMPPEAETLLCYLHLPWHLELLATYTHCLNCPGGAQALQALAHRCPPPL 966

Qy 61 AVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRRTGV 120

Db 967 AVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRRTGV 1026

Qy 121 LVFSELAFLHRSFGDLTAEKQICDLYGRVQARERQALRLRTTFOAFHSAVPFSCG 180

Db 1027 LVFSELAFLHRSFGDLTAEKQICDLYGRVQARERQALRLRTTFOAFHSAVPFSCG 1086

Qy 181 PCLEQQDEERSTRLKDLGRYFEEEGQEPGGMEDAQGPFGQARLQDWDQVRCDIRQF 240

Db 1087 PCLEQQDEERSTRLKDLGRYFEEEGQEPGGMEDAQGPFGQARLQDWDQVRCDIRQF 1146

Qy 241 LSLRPEKFSRAVARIFPHGIGSPCYPAQVYGQDRRFRWKYHLHLSFHALVGLATEELLQV 300

Db 1147 LSLRPEKFSRAVARIFPHGIGSPCYPAQVYGQDRRFRWKYHLHLSFHALVGLATEELLQV 1206

Qy 301 AR 302

Db 1207 AR 1208

RESULT 7

ADU80701

ID ADU80701 standard; protein; 1208 AA.

AC ADU80701;

XX 10-FEB-2005 (first entry)

XX Human RTS-associated protein SEQ ID NO 73.

XX anticancer; apoptosis; RecQ; DNA helicase; RNA interference; RNAi;

KW cytostatic; gene silencing; enzyme.

OS Homo sapiens.

XX WO2004100990-A1.

PN 25-NOV-2004.

XX 19-MAY-2004; 2004WO-JP007145.

XX 19-MAY-2003; 2003JP-00140685.

XX (GENE-) GENE CARE RES INST CO LTD.

PI Takagi M, Shimamoto A, Furuichi Y, Sato A;

XX WPI; 2004-833899/82.

DR Apoptosis inducing agent of cancer cell useful as anticancer agent,

PT comprises compound that suppresses expression of RecQ DNA helicase-family

PT gene/function of protein encoded by RecQ DNA helicase-family gene, as

PT active ingredient.

XX Disclosure; SEQ ID NO 73; 262pp; Japanese.

PS This invention describes a novel anticancer apoptosis inducing agent

CC which comprises (a) compound, which suppresses the expression of RecQ DNA

CC helicase-family gene, (b) DNA which expresses the double-stranded RNA

CC with RNA interfering (RNAi) effect with respect to RecQ DNA helicase-

CC family gene, or (c) a compound, which suppresses the function of protein

CC encoded by RecQ DNA helicase-family gene, as an active ingredient. The

CC invention also describes a method for screening for a candidate compound

CC with a protein encoded by RecQ DNA helicase-family gene. A compound,

CC which suppresses the expression of RecQ DNA helicase-family gene is a

CC transcriptions product or antisense nucleic acid of one part of RecQ DNA

CC helicase-family gene, or a nucleic acid which has ribozyme activity which

CC cleaves the transcription product of RecQ DNA helicase-family gene

CC specifically. It can also be a RecQ DNA helicase-family protein variant

CC which has dominant negative character with respect to the protein encoded

CC by RecQ DNA helicase-family gene, an antibody or a low molecular compound

CC coupled with the protein encoded by RecQ DNA helicase-family gene. The

CC RecQ DNA helicase-family gene is a WRN gene, BLM gene or RecQ1 gene. The

CC method of the invention is useful for manufacturing an anticancer agent

CC as a pharmaceutical composition. The novel agent is highly safe,

CC effective and does not cause any side effects when used as a

CC pharmaceutical. The products of the invention have cytostatic activity.

CC This sequence represents a human helicase-associated RTS protein.

XX Sequence 1208 AA;

SQ Query Match 100.0%; Score 1617; DB 8; Length 1208;

Best Local Similarity 100.0%; Pred. No. 1.2e-158;

Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLTVALDMPPEAETLLCYLHLPWHLELLATYTHCLNCPGGAQALQALAHRCPPPL 60

Db 907 QLTVALDMPPEAETLLCYLHLPWHLELLATYTHCLNCPGGAQALQALAHRCPPPL 966

Qy 61 AVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRRTGV 120

Db 967 AVCLAQQLPEDPGQSSSVFDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRRTGV 1026

Qy 121 LVFSELAFLHRSFGDLTAEKQICDLYGRVQARERQALRLRTTFOAFHSAVPFSCG 180

Db 1027 LVFSELAFLHRSFGDLTAEKQICDLYGRVQARERQALRLRTTFOAFHSAVPFSCG 1086

Qy 181 PCLEQQDEERSTRLKDLGRYFEEEGQEPGGMEDAQGPFGQARLQDWDQVRCDIRQF 240

Db 1087 PCLEQQDEERSTRLKDLGRYFEEEGQEPGGMEDAQGPFGQARLQDWDQVRCDIRQF 1146

Qy 241 LSLRPEKFSRAVARIFPHGIGSPCYPAQVYGQDRRFRWKYHLHLSFHALVGLATEELLQV 300

Db 1147 LSLRPEKFSRAVARIFPHGIGSPCYPAQVYGQDRRFRWKYHLHLSFHALVGLATEELLQV 1206

Qy 301 AR 302

Db 1207 AR 1208

RESULT 8

ADM90956

ID ADM90956 standard; protein; 1306 AA.

AC ADM90956;

XX 03-JUN-2004 (first entry)

XX Human pharmaceutically useful protein SeqID 349.

DE human; cancer; haematopoiesis; thrombosis; anaemia;

XX cardiovascular disorder; ischaemic heart disease;

KW acute myocardial infarction; respiratory disease; asthma; pneumonia;

KW cystic fibrosis; chronic renal failure; glomerulopathy;

KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;

KW HIV infection; systemic lupus erythematosus; endocrine system;

KW diabetes mellitus; epilepsy; Alzheimer's disease;

KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;

KW fungal; parasitic; viral infection; cytostatic; anticoagulant;

KW thrombolytic; antianemic; cardiant; vasotropic; antiasthmatic;

KW antiinflammatory; nephrotropic; antiulcer; hepatotropic;

KW immunosuppressive; antiallergic; dermatological; antirheumatic;

KW antiarthritic; antidiabetic; anticonvulsant; neuroprotective; nootropic;

KW antiparasitic; antibacterial; fungicide; antiparasitic; virucidal;

XX gene therapy; vaccine.

XX Homo sapiens.

OS WO2004020595-A2.

XX 11-MAR-2004.

PD 28-AUG-2003; 2003WO-US027107.

XX PF

XX 29-AUG-2002; 2002US-0406576P.
 PR 29-AUG-2002; 2002US-0406611P.
 PR 29-AUG-2002; 2002US-0406612P.
 PR 29-AUG-2002; 2002US-0406616P.
 PR 29-AUG-2002; 2002US-0406640P.
 PR 29-AUG-2002; 2002US-0406655P.
 PR 29-AUG-2002; 2002US-0406666P.
 PR 17-SEP-2002; 2002US-0410946P.
 PR 17-SEP-2002; 2002US-0410951P.
 PR 17-SEP-2002; 2002US-0410953P.
 PR 17-SEP-2002; 2002US-0410957P.
 PR 17-SEP-2002; 2002US-0410960P.
 PR 17-SEP-2002; 2002US-0410962P.
 PR 17-SEP-2002; 2002US-0411019P.
 PR 17-SEP-2002; 2002US-0411022P.
 PR 17-SEP-2002; 2002US-0411024P.
 PR 17-SEP-2002; 2002US-0411037P.
 PR 17-SEP-2002; 2002US-0411046P.
 PR 17-SEP-2002; 2002US-0411052P.
 PR 17-SEP-2002; 2002US-0411082P.
 PR 17-SEP-2002; 2002US-0411111P.
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.
 PA (DNAP-) DNAFORM KK.
 XX
 PI Williams LT, Chu K, Lee E, Hestir K;
 XX
 DR WPI; 2004-257410/24.
 DR N-PSDB; ADM90747, ADM91165.
 XX
 PT New human polynucleotides and polypeptides, useful for diagnosing,
 PT preventing and treating proliferative disorders, immune disorders,
 PT cardiovascular disorders, or bacterial, fungal, parasitic and viral
 PT diseases.
 XX
 PS Claim 1; SEQ ID NO 349; 254pp; English.
 XX
 CC This invention relates to novel isolated human polynucleotides and the
 CC encoded proteins thereof. Specifically, it refers to proteases, kinases,
 CC phosphatases, secreted and transmembrane proteins, as well as the derived
 CC peptide fragments, which can be used to develop antibodies and screen for
 CC small molecule agonists and antagonists that can modulate their
 CC activities. The present invention describes polypeptides,
 CC polynucleotides, vectors and host cells useful for diagnosing, preventing
 CC and treating proliferative disorders, e.g. cancer, disorders of
 CC haematopoiesis such as thrombosis and anaemia, cardiovascular disorders,
 CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory
 CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the
 CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,
 CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune
 CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders
 CC of the endocrine system, e.g. diabetes mellitus, central nervous system
 CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral
 CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,
 CC parasitic and viral diseases. Accordingly, they exhibit many various
 CC activities including cytostatic, anticoagulant, thrombolytic,
 CC antianaemic, cardiant vasotropic, antiaasthmatic, antiinflammatory,
 CC nephrotropic, antiulcer, hepatotropic, immunosuppressive, anti allergic,
 CC dermatological, antirheumatic, antiarthritic, antidiabetic,
 CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,
 CC fungicide, antiparasitic and virucidal, such that these polynucleotides
 CC can be used for gene therapy purposes and the development of appropriate
 CC vaccines. This polypeptide is a human protein of the invention.
 XX
 SQ Sequence 1306 AA;
 Query Match 98.6%; Score 1594.5; DB 8; Length 1306;
 Best Local Similarity 92.4%; Pred. No. 2.9e-156;
 Matches 302; Conservative 0; Mismatches 0; Indels 25; Gaps 1;
 QY 1 QLTVALDMPPEE-----AIEITLCYLELHPHMLLEIATT 35

Db 980 QLTVALDMPPEEGBEFGVSHRGVGLSPRLSPALPPAIEITLCYLELHPHMLLEIATT 1039
 QY 36 YTHCRINCPGPAQALAHRCPPAVCLAAQQLPEDPGQSSSVFEDMVKLVDSMGWELA 95
 Db 1040 YTHCRINCPGPAQALAHRCPPAVCLAAQQLPEDPGQSSSVFEDMVKLVDSMGWELA 1099
 QY 96 SVRRALCOLQWDHEPRTGVRRGTGVLVFESELAFLHRSFGDLTAEEKDQICDFLYGRVQA 155
 Db 1100 SVRRALCOLQWDHEPRTGVRRGTGVLVFESELAFLHRSFGDLTAEEKDQICDFLYGRVQA 1159
 QY 156 RERQALARLRTPQAFHSVAFPPSCGCLCQDDEERSTRLLKDLGRYFEEBEGEPGGMED 215
 Db 1160 RERQALARLRTPQAFHSVAFPPSCGCLCQDDEERSTRLLKDLGRYFEEBEGEPGGMED 1219
 QY 216 AQGPFGQARLQDWDQVRCDIQFSLRPEEKFSRAVARIFHGIGSPCYPAQVYGQDR 275
 Db 1220 AQGPFGQARLQDWDQVRCDIQFSLRPEEKFSRAVARIFHGIGSPCYPAQVYGQDR 1279
 QY 276 RFWRYLHLSFHALVGLATEELIQQVAR 302
 Db 1280 RFWRYLHLSFHALVGLATEELIQQVAR 1306
 RESULT 9
 ADM44021
 ID ADM44021 standard; protein; 1216 AA.
 XX
 AC ADM44021;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE Mouse RECQL4 protein.
 XX
 KW RECQL4; helicase; protein deactivation; growth disorder;
 KW Rothmund-Thomson syndrome; skin disorder; bone disease;
 KW musculokeletal disease; enzyme.
 XX
 OS Mus musculus.
 XX
 PN WO2005001085-A1.
 XX
 PD 06-JAN-2005.
 XX
 PF 25-JUN-2004; 2004WO-JP009380.
 XX
 PR 27-JUN-2003; 2003JP-00185409.
 XX
 PA (AGEN) NAT INST RADIOLOGICAL SCI.
 XX
 PI Abe M;
 DR WPI; 2005-081681/09.
 DR N-PSDB; ADM44020.
 XX
 PT Novel non-human mammal, preferably rodent such as mouse exhibiting
 PT characteristics of Rothmund-Thomson syndrome, being RECQL4-gene lacking
 PT mouse having mutation in RECQL4-gene, useful as model of human Rothmund-
 PT Thomson syndrome.
 XX
 PS Disclosure; SEQ ID NO 2; 69pp; Japanese.
 XX
 CC This invention describes a novel knock-out mouse which lacks the RECQL4
 CC gene or has a mutation in one of exons 13 to 22 of the RECQL4 gene
 CC resulting in a loss of helicase activity. The mouse described in the
 CC invention exhibits the characteristics of Rothmund-Thomson syndrome and
 CC is suitable as a model of human Rothmund-Thomson syndrome. The mouse
 CC exhibits growth delay, skin abnormality and bone-formation defects
 CC similar to human Rothmund-Thomson syndrome. This sequence represents
 CC murine RECQL4 which is missing in mice exhibiting Rothmund-Thomson
 CC syndrome.
 XX
 SQ Sequence 1216 AA;

Query Match 70.8%; Score 1144.5; DB 9; Length 1216;
 Best Local Similarity 73.2%; Pred. No. 2.1e-109;
 Matches 218; Conservative 26; Mismatches 53; Indels 1; Gaps 1;

QY 1 QLTVALDMPERAEITLLCYLELHPHWHLELLATTYTHCRNLNCPGPAQLQALAHRCPPPL 60
 DB 916 QSTVALDMPERAEITLLCYLELHPHWHLELLPWTYAOCHLHCLGGSQAQLQALAHRCPPPL 975

QY 61 AVCLAQQLPEDPGQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRRTGV 120
 DB 976 AACQAKWPKDTSQGRSSLEFGVVELADSMGKLASVQALHQLKWDPEPKGAAGQGTGV 1035

QY 121 LVFSELAFHLRSPDLTAEBKQICDFLYGRVQARERQALRLRTFOAFHSVAFPSCG 180
 DB 1036 LVKFSSELAFHLRSGDLTDEEKQICDFLYRNVQAREHKAHLHOMSKAFRSVAFPSCG 1095

QY 181 PCLEQDDEERSTRKDLALGRYFEREEGPGMGEDAQGPQCARLQDWEQVRCDIQRF 240
 DB 1096 PCLEQDDEERSTRKDLALGRYFEREEGPGMGEDAQGPQCARLQDWEQVRCDIQRF 1154

QY 241 LSLRPEEFSSRAVARIFHIGISPCYPAQVYQDQRRFWRKYLHLSFHALVGLATEELL 298
 DB 1155 LSLRPEEFSSRAVARIFHIGIASPCYPAQVYGLDRFWRKYLHLDLFHALMHLATEELL 1212

RESULT 10
 ABB63317
 ID ABB63317 standard; protein; 1579 AA.
 AC ABB63317;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 16743.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 XX 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR
 XX N-PSDB; ABL07420.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 XX Disclosure; SEQ ID NO 16743; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1579 AA;
 Query Match 26.4%; Score 427; DB 4; Length 1579;
 Best Local Similarity 31.1%; Pred. No. 1.8e-34;
 Matches 95; Conservative 62; Mismatches 122; Indels 26; Gaps 6;

QY 3 TVQALDMPERAEITLLCYLELHPHWHLELLATTYTHCRNLNCPGPAQLQALAHRCPPPLAV 62
 DB 1290 TVEMLDIPAEINISTLLCTMBLDPFWCISVLSAYVMAKVISYGGPKYLKHAKECPPLAM 1349

QY 63 CLAQQLPEDP-GQSSSVFEDMVKLVDSMGWELASVRRALCOLQWDHEPRTGVRRTGV 121
 DB 1350 AIALQIRDKTFKEDSNIEFSVTDDIAAGIWNWSGVVYQLKLEW--VKVNGYFKRSPIT 1407

QY 122 VEFSELAFHLRSPDLTAEBKQICDFLYGRVQARERQALRLRTFOAFHSVAFPSCGP 181
 DB 1408 VSFYDLGFRKIVPGDFTTESEIDNALDITLTSVKQERTQLIQYVAHGLAAVAYSSCGQ 1467

QY 182 CLEQQ-DEERSTRKDLALGRYFEREEGQ-----EPGGMEDAQGPQCARLQDWEQV 234
 DB 1468 CCNADFFQDRGEQKALVRNYFANDYPQDLELELETPSNVPD-----ENII 1512

QY 235 CDIRQFLSLRPEEFSSRAVARIFHIGISPCYPAQVYQDQRRFWRKYLHLSFHALVGLAT 294
 DB 1513 DDVHALINMYPDNTFTGTNTARIIFHIGIMSPNYPAVINGR-CKFWRAHVKVDFNRILHLAN 1571

QY 295 EELQ 299
 DB 1572 MAIK 1576

RESULT 11
 ADY13119
 ID ADY13119 standard; protein; 206 AA.
 XX
 AC ADY13119;
 XX
 DT 21-APR-2005 (first entry)
 DE
 XX Plant full length insert polypeptide seqid 68934.
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.
 XX
 OS Unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 XX 28-APR-2003; 2003US-00425114.
 PF
 XX 06-MAY-1999; 99US-00304517.
 PR
 XX 05-NOV-2001; 2001US-00985678.
 XX
 XX (LIUJ/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABAV/) TABASKA J E.
 XX (CAOY/) CAO Y.
 XX
 PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 XX WPI; 2004-180133/17.
 DR
 XX

PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.

PS Claim 1; SEQ ID NO 68934; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp://seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.

XX SQ Sequence 206 AA;

Query Match 9.0%; Score 145.5; DB 8; Length 206;
Best Local Similarity 25.0%; Pred. No. 2.4e-06;
Matches 50; Conservative 32; Mismatches 69; Indels 49; Gaps 9;
QY 123 EFS-ELAFHLRSPG-----DLTAEKDKQICDFLYGRVQARERQALRLRTQAF 171
DB 29 QFSGEISFELKPCACIYVILKKPDDLNALSAD-----ITRWLSEVNSKSKLDAMFD-L 82
QY 172 HSAVFPSC-----GPCLEQDEERSLTKDLLGRYFEEEGQEPGGMEDAQGPE 220
DB 83 ANFAVKGCQRTDGCSGSQTPTCIQK-----IIQVFSKNYSTSDSG--QCTOPQ 129
QY 221 PGQARLQDQWEDVRCDDIRQFLSLRPEKFSRAVARIFHGISPCYPAQVYQDQRRWRK 280
DB 130 RGSFPLQ-----ADIKVFLKNSFAKPTPRAVARIMHGISSPAFFSATWSKN-HFWGR 181
QY 281 YLHLSHFALVGLATEELLQV 300
DB 182 YLEVDFFVMEAKAELVKL 201

RESULT 12
ABO80677

ID ABO80677 standard; protein; 762 AA.

AC ABO80677;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #12852.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

FN US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD14248.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 29423; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67836-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 762 AA;

Query Match 6.7%; Score 108.5; DB 7; Length 762;
Best Local Similarity 23.9%; Pred. No. 0.1;
Matches 72; Conservative 21; Mismatches 79; Indels 129; Gaps 14;
QY 46 GPAQLQALAH---RCPPLAVCLAQQLPEDPGSGSSVFEFDMVKLVDSMGWELASVRRALC 102
DB 470 GPSLGGPRLRRPRRRPPLA-----ETPGQRRSAEH-----APGTGSD----- 507
QY 103 QLQWDHPRGTGVRGTGVLVEFSELAFLHRSFGDLTAEKDKQICDFLYGRVQARERQA-- 160
DB 508 ---DHGPRDPGR-----DRPGELVQSRPRRSCAAGQGVHRRLRQADG 547
QY 161 -----LRLRRTFOAFHSVAFPPCGPCLEQDEERSTRLKDLLGRYFEEEGQEP 210
DB 548 GRPRATAGRLPLRLR-----RPGFHLPSRTER--RLAQVLGRWRDEPRAP 594
QY 211 GGMEDA-----QGP-----EPQARLQDQWEDQV 233
DB 595 GGHRLARRDGHPRRAQRAAMGFLRCVPQHQQGPGDQARGQAALQLPPGLRLDQKRAV 654
QY 234 RCDIRQFLSLRPEEKFSRAVARIFHG-----IGSPCY----PAQVYQDQR 275
DB 655 QDQLPRRIP-RADRRRRRRAPASAGEGPPATDQPHRRHHPRVGDPHHFGQRPRA---QGRP 710
QY 276 R 276
DB 711 R 711

RESULT 13

ABR82253

ID ABR82253 standard; protein; 751 AA.

AC ABR82253;

XX 13-OCT-2003 (first entry)

XX Human nucleic acid-associated protein (NAAP)-Id 1532441CD1.

XX NAAP; nucleic acid-associated protein; cardiant; cytostatic; transgenic;
XX neuroprotective; gene therapy; human; cancer.

XX

OS Homo sapiens.
XX
XX WO2003052048-A2.
XX
XX 26-JUN-2003.
XX
XX 02-MAY-2002; 2002WO-US014276.
XX
XX 04-MAY-2001; 2001US-0288598P.
XX 17-MAY-2001; 2001US-0291776P.
XX 18-MAY-2001; 2001US-0292172P.
XX 25-MAY-2001; 2001US-0293564P.
XX
XX (INCYTE GENOMICS INC.
XX
XX Yue H, Ding L, Baughn MR, Lal PG, Yue H, Hafalia AJA, Lee EA,
XX Ison CH, Becha SD, Gururajan R, Emerling BM, Griffin JA, Tang YT,
XX Lu DAM, Yao MG, Chawla NK, Ramkumar J, Gandhi AR, Lee SY,
XX Richardson TW, Yang J, Elliott VS, Lu Y, Thangaveelu K, He A,
XX Azimzal Y, Raumann BE, Swarnakar A, Burford N;
XX
XX WPI; 2003-541640/51.
XX
XX N-PSDB; ACF355661.
XX
XX New human nucleic acid-associated proteins polypeptide, useful for
XX preparing a composition for diagnosing or treating e.g., cardiovascular
XX or neurological disorders.
XX
XX Claim 1; Page 185-187; 212pp; English.
XX
XX The invention relates to human nucleic acid-associated proteins (NAP)
XX and encoding polynucleotides. The NAP polypeptides can be expressed by
XX standard recombinant methodology. The polypeptides are useful for
XX preparing a composition for diagnosing or treating a disease or condition
XX associated with decreased expression or overexpression of functional NAP
XX e.g., cardiovascular or neurological disorders or cancer. The present
XX sequence represents a human NAP polypeptide
XX
XX Sequence 751 AA;
XX

Query Match 6.6%; Score 107; DB 6; Length 751;
Best Local Similarity 27.0%; Pred. No. 0.14;
Matches 70; Conservative 23; Mismatches 106; Indels 60; Gaps 10;
QY 11 BEAETILCYLELHPHMHLELATTY-----HCRINCPCGPAQALAHRCPLAVCLA 65
DB 159 BEARRRQAAQCPQHGEALRFLCQPCSQLCRECHLD-----PHLDHPLPLAEAVR 210
QY 66 QQLPEDPGQSSSVEFDMVKLVDSMGWELASVVR-----ALCOLQMD-----HEPTG 113
DB 211 ARPPGLEG-----LLAGVDNNLVELEARRVKEKALARLQQAARVGTQVEAAAG 261
QY 114 VRRGTGVLFESBELAHLRSPGDLTAEBKQICDPLYGVO-ARERQALAR----- 163
DB 262 VLR--ALTAQKQEVIGQLRAHVAEAAERLAELEGREQVAAAAAFARRVLSIGREA 319
QY 164 -----LRRTPQAFHSVAFSPGCPCLAEQDDEESTRLK---DLGGYFEEBEQEPG 211
DB 320 EILSLGALIAORLRQLOQCPWAGPAPCILLPQLEHPLGLDKXCHLLRSLFEBQDPQKXG 379
QY 212 GMEIDA--QGPEQOARLQD 228
DB 380 GKDGAGTQGGESQSHRED 398

RESULT 14
ABM80914
ID ABM80914 standard; protein; 755 AA.
XX
XX ABM80914;
XX
XX 18-NOV-2004 (first entry)
XX

DE Tumour-associated antigenic target (TAT) polypeptide PRO81867, SEQ:2360.
XX
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
XX tumour; diagnosis; cell proliferative disorder; breast cancer;
XX colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX central nervous system cancer; bladder cancer; pancreatic cancer;
XX cervical cancer; melanoma; leukaemia; hybridisation probe;
XX chromosome identification; chromosome mapping; gene mapping;
XX gene therapy; cytostatic.
XX
XX Homo sapiens.
XX
XX WO2004030615-A2.
XX
XX 15-APR-2004.
XX
XX 29-SEP-2003; 2003WO-US028547.
XX
XX 02-OCT-2002; 2002US-0414971P.
XX
XX (GETH) GENENTECH INC.
XX
XX Mu TD, Zhang Z, Zhou Y;
XX
XX WPI; 2004-347921/32.
XX
XX N-PSDB; ACN38689.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
XX useful in preparing a medicament for treating or detecting a
XX proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX prostate cancer or tumor.
XX
XX Claim 12; SEQ ID NO 2360; 7273pp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are
XX overexpressed in cancer tissues compared to normal tissues, and may thus
XX serve as effective targets for the diagnosis and treatment of cancer in
XX mammals. The invention also relates to nucleic acid and polypeptide
XX sequences at least 80% identical to the TAT nucleic acid and
XX polypeptides; expression vectors and host cells comprising a TAT nucleic
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX TAT polypeptide; and methods and compositions for the treatment or
XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX antibodies, antagonists, binding molecules and compositions are useful
XX for diagnosing or treating a cell proliferative disorder associated with
XX increased TAT expression, particularly cancers such as breast cancer,
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX cancer, pancreatic cancer, cervical cancer, cancers of the central
XX nervous system, melanoma and leukaemia. TAT nucleic acids may further be
XX used as hybridisation probes, in chromosome and gene mapping, in
XX chromosome identification and in gene therapy. The present sequence
XX represents a TAT polypeptide of the invention
XX
XX Sequence 755 AA;
XX

Query Match 6.6%; Score 107; DB 8; Length 755;
Best Local Similarity 27.0%; Pred. No. 0.15;
Matches 70; Conservative 23; Mismatches 106; Indels 60; Gaps 10;
QY 11 BEAETILCYLELHPHMHLELATTY-----HCRINCPCGPAQALAHRCPLAVCLA 65
DB 159 BEARRRQAAQCPQHGEALRFLCQPCSQLCRECHLD-----PHLDHPLPLAEAVR 210
QY 66 QQLPEDPGQSSSVEFDMVKLVDSMGWELASVVR-----ALCOLQMD-----HEPTG 113
DB 211 ARPPGLEG-----LLAGVDNNLVELEARRVKEKALARLQQAARVGTQVEAAAG 261
QY 114 VRRGTGVLFESBELAHLRSPGDLTAEBKQICDPLYGVO-ARERQALAR----- 163
DB 262 VLR--ALTAQKQEVIGQLRAHVAEAAERLAELEGREQVAAAAAFARRVLSIGREA 319

```
QY 164 -----LRTFOAFHSVAFPSGPGCLBQDERSTRK---DLGRYFEERGGSPG 211
Db 320 EILSLKGLAQLRLQGGCPMAFPAPCLLPQLLEHPLGLDKNCHLRLSFEEDQPPQKDG 379
QY 212 GMEBA--OGPSPGQARLQD 228
Db 380 GKDAGTGGGSESGORRED 398

RESULT 15
AA96736
ID AA96736 standard; protein; 1029 AA.
AC AA96736;
DT 26-SEP-2000 (first entry)
DE PRO3434, a novel secreted protein.
XX PRO3434; secreted protein; transmembrane protein; recombinant production;
KM gene therapy.
XX Homo sapiens.
XX Key
XX Peptide
XX 1..16
XX /label= Signal_peptide
XX 91..97
XX /note= "N-myristoylation site"
XX 136..142
XX /note= "N-myristoylation site"
XX 154..158
XX /note= "cAMP- and cGMP-dependent protein kinase
XX phosphorylation site"
XX 224..230
XX /note= "N-myristoylation site"
XX 329..333
XX /note= "amidation site"
XX 331..335
XX /note= "cAMP- and cGMP-dependent protein kinase
XX phosphorylation site"
XX 435..441
XX /note= "N-myristoylation site"
XX 439..445
XX /note= "N-myristoylation site"
XX 443..449
XX /note= "N-myristoylation site"
XX 616..620
XX /note= "cAMP- and cGMP-dependent protein kinase
XX phosphorylation site"
XX 634..638
XX /note= "amidation site"
XX 665..671
XX /note= "N-myristoylation site"
XX 698..704
XX /note= "N-myristoylation site"
XX 785..789
XX /note= "cAMP- and cGMP-dependent protein kinase
XX phosphorylation site"
XX 891..895
XX /note= "cAMP- and cGMP-dependent protein kinase
XX phosphorylation site"
XX W0200036102-A2.
XX 22-JUN-2000.
XX 01-DEC-1999; 99WO-US028634.
XX 16-DEC-1998; 98US-0112851P.
XX 16-DEC-1998; 98US-0113145P.
XX 22-DEC-1998; 98US-0113511P.
XX 12-JAN-1999; 99US-0115556P.
```

```
PR 12-JAN-1999; 99US-0115556P.
PR 12-JAN-1999; 99US-0115733P.
PR 09-FEB-1999; 99US-0119341P.
PR 10-FEB-1999; 99US-0119537P.
PR 12-FEB-1999; 99US-0119565P.
PR 02-JUN-1999; 99WO-US012252.
XX (GETH ) GENENTECH INC.
XX Botstein D, Desnovers L, Ferrara N, Fong S, Gao W, Goddard A,
PI Gurney AL, Pan J, Roy MA, Stewart TA, Tumas D, Watanabe CK,
PI Wood WJ;
XX WPI; 2000-431586/37.
DR N-PSDB; AAA51266.
XX Isolated nucleic acid molecule encodes a PRO polypeptide which is a
PT transmembrane polypeptide.
XX Claim 12, Fig 16; 154pp; English.
XX This is PRO3434, a novel secreted protein. The invention concerns novel
CC secreted and transmembrane proteins, designated PRO polypeptides. The
CC cDNA and gene sequences are useful in the recombinant production of PRO
CC polypeptides, as a hybridization probe to screen libraries to isolate
CC cDNAs with sequence identity to PRO polypeptides or to map the gene
CC encoding the PRO polypeptides and analyzing genetic disorders. The
CC cDNA/gene can also be used to produce transgenic animals useful for the
CC development and screening of therapeutically useful reagents. They can
CC also be used in gene therapy, e.g. to replace a defective gene
XX Sequence 1029 AA;
```

Query Match 6.6%; Score 107; DB 3; Length 1029;
Best Local Similarity 26.2%; Pred. No. 0.22;
Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

```
QY 29 LELLATYYTHGRNLNCPGPPAQLALHRCPLAVCLAQQL-----PEDPGGS--SSV 79
Db 245 LQALATL-----LSPFGGALVMSM-HRSHFLACPRLRLQCYORCVPODTGFSLEFLKV 298
QY 80 EFDWVKLVDSWGMGLASVRBALCOL--QMDHEPRTGVRGTVVERSELAFH----- 130
Db 299 LLQMLQWLDSPGVGCPPLRAQLKMLASQASAGRLSPVRG-GLIRLAEDALFRODLEVS 357
QY 131 -----LSPGDLTAEEKQICDFLYGRVQARERQALARLRTFQAHSVA--FPS 179
Db 358 STVRAVATATRS-GEQGSVEPDLISKVLQGLIVRSPHLEILTAFFSATADASPFPAC 416
QY 180 GPCLBQDERSTRKDLGRYFEER--QGP--GMEBAQSPPGQARLQDWEDQV 233
Db 417 KPVVV-----VSSIL--LQEEPLAGKKGAGGGSLEAVRLPSSGLVDWLEML 464
QY 234 RCDIRQFLSLRPERK-----FSSRAVARIFHGIGSPCPAQQVGGDRRFWRKY----- 281
Db 465 DPEV---VSSCPDLQLRLFSRR-----KKGGAQVPS-----FRPYLLTLFTHQ 506
QY 282 -----LHLSFHALVGLATEE 296
Db 507 SSWPTLHQICIRVLGKSRQ 526
```

Search completed: December 27, 2005, 21:48:09
Job time : 57.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:40:26 ; Search time 11.4 Seconds

(without alignments)
2548.900 Million cell updates/sec

Title: US-09-889-325-4_COPY_907_1208

Perfect score: 1617
Sequence: 1 QLTQALDMPREAIETLLCY.....HLSPHALVGLATEELLQVAR 302Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summariesDatabase :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|---------------------|
| 1 | 186 | 11.5 | 941 | 2 | A86404 | probable protein A |
| 2 | 97.5 | 6.0 | 1548 | 2 | T25808 | hypothetical prote |
| 3 | 96.5 | 6.0 | 976 | 2 | S40697 | processing endopro |
| 4 | 95 | 5.9 | 1582 | 2 | E70876 | probable polyketid |
| 5 | 94.5 | 5.8 | 298 | 2 | TC7568 | kidney inhibitor o |
| 6 | 91.5 | 5.7 | 1397 | 2 | T46354 | hypothetical prote |
| 7 | 91.5 | 5.7 | 2548 | 2 | B59435 | myosin IIA [import |
| 8 | 89 | 5.5 | 2274 | 2 | TC0258 | adenomatous polypo |
| 9 | 88.5 | 5.5 | 308 | 2 | JC2549 | apolipoprotein F p |
| 10 | 87 | 5.4 | 679 | 2 | E71282 | probable NH(3)-dep |
| 11 | 85.5 | 5.3 | 548 | 2 | C75499 | cytidine/deoxycyt |
| 12 | 85.5 | 5.3 | 634 | 2 | T00359 | hypothetical prote |
| 13 | 85 | 5.3 | 637 | 2 | B95878 | probable adenylyat |
| 14 | 85 | 5.3 | 637 | 2 | JC7753 | ring finger B-box |
| 15 | 84.5 | 5.2 | 476 | 2 | T42692 | hypothetical prote |
| 16 | 84.5 | 5.2 | 1209 | 2 | T00373 | hypothetical prote |
| 17 | 84.5 | 5.2 | 4687 | 1 | A39638 | plectin - rat |
| 18 | 84 | 5.2 | 472 | 2 | IS9087 | ISG-K54 - human |
| 19 | 84 | 5.2 | 1175 | 2 | IS7549 | adenosine deaminas |
| 20 | 83.5 | 5.2 | 343 | 2 | HT0512 | probable LPD prot |
| 21 | 83.5 | 5.2 | 574 | 2 | B57072 | hypothetical prote |
| 22 | 83 | 5.1 | 547 | 2 | T45635 | hypothetical prote |
| 23 | 83 | 5.1 | 948 | 2 | T03225 | probable regulator |
| 24 | 83 | 5.1 | 1082 | 2 | T13410 | hypothetical prote |
| 25 | 82.5 | 5.1 | 585 | 2 | S49253 | rubisco binding pr |
| 26 | 82.5 | 5.1 | 1097 | 2 | G85092 | hypothetical prote |
| 27 | 82.5 | 5.1 | 1157 | 2 | H69163 | DNA helicase relat |
| 28 | 82.5 | 5.1 | 1545 | 2 | T14288 | DNA (cytosine-5-)- |
| 29 | 82 | 5.1 | 1118 | 2 | C95385 | probable adenylyate |

| | | | | | | |
|----|------|-----|------|---|--------|--------------------|
| 30 | 81.5 | 5.0 | 1211 | 2 | C83110 | probable exonuclea |
| 31 | 81.5 | 5.0 | 2626 | 2 | T31099 | myosin-RhoGAP prot |
| 32 | 81.5 | 5.0 | 4574 | 2 | G02520 | plectin - human |
| 33 | 81.5 | 5.0 | 4684 | 2 | A59404 | plectin [imported] |
| 34 | 81 | 5.0 | 440 | 1 | BVBRCD | cyad protein - Bor |
| 35 | 81 | 5.0 | 473 | 2 | A48949 | beta-glucosidase, |
| 36 | 80.5 | 5.0 | 537 | 2 | B87414 | RNA polymerase sig |
| 37 | 80.5 | 5.0 | 1013 | 2 | S32214 | hypothetical prote |
| 38 | 80.5 | 5.0 | 1689 | 2 | S72467 | sodium channel pro |
| 39 | 80 | 4.9 | 507 | 2 | T36009 | probable glucose-6 |
| 40 | 80 | 4.9 | 1148 | 2 | AD0198 | transcription-repa |
| 41 | 80 | 4.9 | 1334 | 2 | B86451 | probable copia-tyr |
| 42 | 79.5 | 4.9 | 242 | 1 | RHHUT | thyroliberin precu |
| 43 | 79.5 | 4.9 | 293 | 1 | QOCVPY | B1 protein - pota |
| 44 | 79.5 | 4.9 | 456 | 2 | T39014 | hypothetical zinc- |
| 45 | 79.5 | 4.9 | 747 | 2 | I39444 | AMP deaminase (EC |

ALIGNMENTS

RESULT 1

A86404
probable protein ATP-dependent DNA helicase RecQ [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86404
R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Martelli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-941 <STO>
A:Cross-references: UNIPARC:UPI000017A6AC; GB:AE005172; NID:g10998928; PIDN:AG26068.1; C:Genetic:

A:Map position: 1

Query Match 11.5%; Score 186; DB 2; Length 941;

Best Local Similarity 24.4%; Pred. No. 1e-07;
Matches 77; Conservative 45; Mismatches 126; Indels 68; Gaps 12;

| | | | |
|----|-----|--|----------|
| QY | 5 | QALDMPREAIETLLCYELPHHMLLELATYTHGRINCPGPAQLQALNR--- | CPPLA 61 |
| DB | 659 | QFDDKKEEWMQITLLHLEGEVOYLRLPOLNICCLN-----FHRTSPAQNTIV 707 | |
| QY | 62 | VCIAOOLPEPDQSGSSVFDVWKLVDSWG--WELASYRALCOLQMDHEBRTGVRGT 118 | |
| DB | 708 | MLYQSSPTTLAARAIYVAALKSHVKGGLHVPDIPAVASSICVAT-----T 755 | |
| QY | 119 | GVLVSELSAFLURSPGDLTAEEKD-----QICDFLVGRVQAREFOALRLR 166 | |
| DB | 756 | DVLAETQAL-----KGEVTEYELKDSAFCYTIKSPKEICSLSHLTWLTETIESCKVRK 809 | |
| QY | 167 | -TFQAFHSVAFSPCCPCBCQODEEESTRLKDLGVEEESBQBPQGMDAQEPGQAR 225 | |
| DB | 810 | LIMMSAAVAALSVSNTSLSSGAKQTR--SLQSRIFDFNG-----DEKCDSPSKA- 859 | |
| QY | 226 | LDQWEDQ---VRCDIROFLSLRPEKSSRAVARIFGIGSPCYPAQYVQDRFRWKY 281 | |
| DB | 860 | -----TONCAFPRADIKVFLQGNRQAKFPPRAIIMHGVGSPAPFNSVWSK-TTFWGRY 913 | |
| QY | 282 | LHLSPHALVGLATEEL 297 | |
| DB | 914 | MNVDFRIVMEAAQTEL 929 | |

RESULT 2

T25808

C/Specties: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 05-Oct-2004

C/Accession: T25808

R/Becker, M.; Mohlmann, P.

submitted to the EMBL Data Library, April 1997

A/Description: The sequence of C. elegans cosmid K08B12.

A/Reference number: Z20091

A/Accession: T25808

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1548 <BEC>

A/Cross-references: UNIPROT:O01583; UNIPARC:UP1000017A3A8; EMBL:U97001; PIDN:AAB52260.1;

A/Experimental source: strain Bristol N2; clone K08B12

C/Genetics:

A/Gene: CESP:K08B12.5

A/Map position: 5

A/Intons: 39/1; 66/1; 156/3; 278/2; 391/3; 628/1; 1022/2; 1273/3; 1326/1; 1423/2; 1489/

F/956-1005/Domain: protein kinase C zinc-binding repeat homology <K2N>

Query Match 6.0%; Score 97.5; DB 2; Length 1548;

Best Local Similarity 24.8%; Pred. No. 6.3;

Matches 60; Conservative 38; Mismatches 103; Indels 41; Gaps 11;

QY 49 QQLAHRCPLAVCLAQQLPEDPGQSSSVFPMVKLVDSMGWELASVRALCOLQMD 107

Db 499 EIQILNKRLEDEALAQOQKPKDEIVASEKKLKEIKERNQVLWKESEIQRELDNIN-D 557

QY 108 HERTGVRCGVVSEBELAFHLRSPBDLVAEKD-----QICPLGRV 153

Db 558 HLDVLVEKAT-VVOQRDMQAEIADVDLSLITKDSVKRLQDEAKKAKQVADF-EKTL 615

QY 154 QAREROALRLR-----TFOAFHS-----VAFPSGPGCLEQDERSTRKDL 198

Db 616 KEITEKIALIKKBEVTEIKRSVEYDDHLSSEVVAANKNTIASIQATNBSRETEIKLK 675

QY 199 GRVFE---EEGQEPGMEDAQP-EPGQARLQDMEDVRCDIRQFLSLRPE-EKFSRA 253

Db 676 QRMEERASHTAQSEQEKQLEAHYERAKQKMLQDNVEGMNENR---GLRDEIKLSQM 732

QY 254 VA 255

Db 733 AA 734

RESULT 3

S40697

C/Specties: Yarrowia lipolytica, Candida lipolytica

C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004

C/Accession: S40697

R/Benderlin, C.S.; Ogrzydzak, D.M.

Yeast 10, 67-79, 1994

A/Title: Cloning, nucleotide sequence and functions of XPR6, which codes for a dibasic p

A/Reference number: S40697; MUID:94262316; PMID:8203153

A/Accession: S40697

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-976 <END>

A/Cross-references: UNIPROT:P42781; UNIPARC:UP1000013901C; EMBL:L16238; NID:q295682; PIR

F/302-542/Domain: subtilisin homology <SP7>

F/311,349,528/Active site: Asp, His, Ser #statue predicted

Query Match 6.0%; Score 96.5; DB 2; Length 976;

Best Local Similarity 21.8%; Pred. No. 4.4;

Matches 56; Conservative 33; Mismatches 91; Indels 77; Gaps 10;

QY 40 RINCPGPAQLALHRCPLAVCLAQQLPEDPGQSSSVFPMVKLVDSMGWELASVR 99

Db 89 RLAVAGVLAKEELRK-----EKLOKKGMSSEDEVEKRLV----- 124

QY 100 ALCOLQMDHBPRTGVRGGTVLVEFSELAFHLSPGDLTAEEKDQICDPLVGRVQAREQ 159

Db 125 ALERLDYDWS-----ERGLASLEVLSERRJHKAPVWVTEEMEYLKE-----IKRAEE 174

QY 160 ALARLRTPQAFHSHVAFPSGPGCLEQDERSTRKQLLGRVPEEERGQEPGMEDAQP 218

Db 175 A-----QKAQDDKGGKKEQDKDKKEGQDEAKQKEDNKGD 211

QY 219 -PEPGARLQDMEDQVRCIRQFLSLRPEKFSRAVARIFHGISP-----CYP 267

Db 212 DKEDGEEDDDDEDDED-----ASPANPVQMKPVDESMTG-GMPDDSLYDYRKRYTP 264

QY 268 AQYVGQDRFRFRK-YLH 283

Db 265 DEVGIRKPSLMKQWYLH 281

RESULT 4

E70876

C/Specties: Mycobacterium tuberculosis (strain H37RV)

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C/Accession: E70876

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: E70876

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1582 <COL>

A/Cross-references: UNIPROT:O50437; UNIPARC:UP1000003ADG; GB:AL010186; GB:AL123456; NID:

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: pke4

C/Superfamily: Mycobacterium tuberculosis probable polyketide synthase pke4; acyl carrier

ology

C/Keywords: carrier protein

F/35-315/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>

F/1242-1422/Domain: short-chain alcohol dehydrogenase homology <SADH>

F/1506-1578/Domain: acyl carrier protein homology <ACP1>

Query Match

5.9%; Score 95; DB 2; Length 1582;

Best Local Similarity 22.2%; Pred. No. 10;

Matches 58; Conservative 29; Mismatches 72; Indels 102; Gaps 13;

QY 44 PGCPAQLALHRCPLAVCLAQQLPED-----PGQSSSVFDM 83

Db 791 PGEPRLPVVTRSA-----ASVLPSDLANLEQAGIRGMRYDSHPHKGATLIDVN 843

QY 84 VKLVDSMGWELASVRALCOLQ-WDHEPRTGVRG-----TGLVEFS 125

Db 844 DEYVAAL---VAS-----QLQSGSQDEDEAWMNGIWTARLRGRLRPAERRTAVVEYR 894

QY 126 E--LAFHLSPGDLTAEEKQICDPLVGRVQAREQALRLRTPQAFHSHVAFPSGPGCL 183

Db 895 RDGMRQLQIRTPGDLSELE-----FVTFDRVA-FGPGE-I 926

QY 184 EQQDERSTRKDL-----GRY-----FEERGOEPGMEDAQEPGQARLQD----- 228

Db 927 EVAVTASSVVFADVLVAFGRYPTFEGTRQQLGIDFAGVTVAVGDVYTHRGHVGMSA 986

QY 229 ---WEDQVRCDIRQFLSLRPE 246

Db 987 NGCMSTFVRCDARLAVTLRPE 1007

RESULT 5

UC7568
Kidney inhibitor of apoptosis protein - human
C/Species: Homo sapiens (man)
C/Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C/Accession: UC7568
R/Lin, J.H.; Deng, G.; Huang, Q.; Morser, J.
Biochem. Biophys. Res. Commun. 279, 820-831, 2000
A/Title: Klap, a novel member of the inhibitor of apoptosis protein family.
A/Reference number: UC7568; MUID: 21092523; PMID:11162435
A/Contents: Fetal Kidney
A/Accession: UC7568
A/Molecule type: mRNA
A/Residues: 1-298 <LIN>
A/Cross-references: UNIPROT:Q96CA5; UNIPARC:UPI00001269A6
C/Comment: This protein, a new member of the inhibitor of apoptosis protein family, plays
C/Genetics:
A/Gene: klap
A/Map position: 20q13.3
C/Keywords: apoptosis

Query Match 5.8%; Score 94.5; DB 2; Length 298;
Best Local Similarity 20.8%; Pred. No. 1.6;
Matches 62; Conservative 32; Mismatches 113; Indels 91; Gaps 12;

QY 25 PNNWELATTYTHGRINCPGPPAQQLAHRCPLAVCLAQQLPEPDQSSSVFPMV 84
DB 16 PSHMA-----AGDEPTQ-----ERCGRPSL-----GSPVGLDTC 45
QY 85 KLVDSMGWELASVRALCOLQMDHEPRTGVRGTGV-----LVFSELAFLHSPG 135
DB 46 RAMDVNDGIIQLRLPTEEBEBGAGATLSRGPAPGSGSEIRLASVYDMPPLTAEPV 105
QY 136 DITAE-----EKDQI-CDPLYGRVQARER-----QALARL-----RT 167
DB 106 ELLAAAGFPHTGHDQKRCFCFCYGLQSWKRGDPWTEHAKVFPSCQFLIRSGKGRDFVHS 165
QY 168 PQAHSVAPPSGCPCLQEODDERSTRKDLG-----RYEEBEGEPGMEDAQSP 219
DB 166 VQETHSQLIGSWDPWEPEPDAAPVAPSVAPSGYBELPTPRREVQSSAQEPGVSPAEAO 225
QY 220 -----EPGQARLQDWEQVRCDIROFLSLRPEKFSRAVARIF-----HGIGSPCY 267
DB 226 RAMWVLEPFGAR-----DVEAQLRLQEBERTCKVCLDAVSLVFPCHLVCAECAP 277

RESULT 6

T46354
hypothetical protein DKFZp434F1016.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T46354
R/Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23037
A/Accession: T46354
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1397 <AAA>
A/Cross-references: UNIPROT:Q9NTG2; UNIPARC:UPI0000071438; EMBL:AL137287
A/Experimental source: adult testis; clone DKFZp434F1016
C/Genetics:
A/Note: DKFZp434F1016.1
F/849-897/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 5.7%; Score 91.5; DB 2; Length 1397;
Best Local Similarity 21.5%; Pred. No. 18;
Matches 55; Conservative 38; Mismatches 76; Indels 87; Gaps 12;
QY 47 PAQLQALAHRCPLAVCLAQQLPEDPG-----QGSSSVFPMVKLVDSM 90
DB 647 PDELAAY-HPTPLS-----PELPGSCRKEFKENKPSPAKRRKRSVKSVAL-DSM 697

QY 91 GWEIASVVRALCOLQMDHEPRTGVRGTGVLFSESL-----AFHLRSPGDLTAEE--KDO 144
DB 698 HMQNDSVQ-----IIASVDSLKMDSEFLTKKYNDLNDSDSKDT 736
QY 145 ICDPLYGRVQARERQALRLRTQAFHSVAFPSGCGCLBEQDERSTRKDLGGRYEE 204
DB 737 LVDPVF-----KKALKEFRQNIFFSYSSAL-----AMDGKSIRYKDLVAFEQI 781
QY 205 BEGEPGMEDAQEPPEQARLQDWEQVRCDIROFLSLRPEKFSRAVARIFHGIGSP 264
DB 782 LEKTRLEQRDSLGSF-----VRVWNTFKVFLDEYNN-----EFKT-----SD 821
QY 265 CYPAGVYGQDRRFWRK 280
DB 822 CTATKVPKTERKGRK 837

RESULT 7

E59435
myosin IXA [imported] - human
C/Species: Homo sapiens (man)
C/Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C/Accession: E59435
R/Gorman, S.W.; Halder, N.B.; Grieshammer, U.; Swiderski, R.E.; Kim, E.; Welch, J.W.; See
Genomics 59, 150-160, 1999
A/Title: The cloning and developmental expression of unconventional myosin IXA (MYO9A) a
A/Reference number: E59435
A/Accession: E59435
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-2548 <GOR>
A/Cross-references: UNIPROT:Q9UNJ2; UNIPARC:UPI00002706F; GB:NP_008832; PTD:95902012; P)

Query Match 5.7%; Score 91.5; DB 2; Length 2548;
Best Local Similarity 21.5%; Pred. No. 37;
Matches 55; Conservative 38; Mismatches 76; Indels 87; Gaps 12;

QY 47 PAQLQALAHRCPLAVCLAQQLPEDPG-----QGSSSVFPMVKLVDSM 90
DB 1798 PDELAAY-HPTPLS-----PELPGSCRKEFKENKPSPAKRRKRSVKSVAL-DSM 1848
QY 91 GWEIASVVRALCOLQMDHEPRTGVRGTGVLFSESL-----AFHLRSPGDLTAEE--KDO 144
DB 1849 HMQNDSVQ-----IIASVDSLKMDSEFLTKKYNDLNDSDSKDT 1887
QY 145 ICDPLYGRVQARERQALRLRTQAFHSVAFPSGCGCLBEQDERSTRKDLGGRYEE 204
DB 1888 LVDPVF-----KKALKEFRQNIFFSYSSAL-----AMDGKSIRYKDLVAFEQI 1932
QY 205 BEGEPGMEDAQEPPEQARLQDWEQVRCDIROFLSLRPEKFSRAVARIFHGIGSP 264
DB 1933 LEKTRLEQRDSLGSF-----VRVWNTFKVFLDEYNN-----EFKT-----SD 1972
QY 265 CYPAGVYGQDRRFWRK 280
DB 1973 CTATKVPKTERKGRK 1988

RESULT 8

T30258
adenomatous polyposis coli protein 2 - mouse
N/Alternate names: APC2 protein
C/Species: Mus musculus (house mouse)
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C/Accession: T30258
R/van Be, J.H.; Kirkpatrick, C.; van de Wetering, M.; Molenaar, M.; Miles, A.; Kuipers, J.
Curr. Biol. 9, 105-108, 1999
A/Title: Identification of APC2, a homologue of the adenomatous polyposis coli tumour sui
A/Reference number: Z20796; MUID:99147086; PMID:10021369
A/Accession: T30258
A/Status: preliminary
A/Molecule type: DNA

A:Residues: 1-2274 <VAN>
A:Cross-references: UNIPROT:Q9Z1K7; UNIPARC:UPI00000296CC; EMBL:AJ130783; NID:g4210431;
C:Genetics:
A:Gene: Apcc2
A:Introns: 47/3; 78/1; 138/2; 174/3; 212/3; 238/3; 271/3; 396/1; 428/1; 474/3; 500/3; 53

Query Match 5.5%; Score 89; DB 2; Length 2274;
Best Local Similarity 21.9%; Pred. No. 52;
Matches 68; Conservative 42; Mismatches 113; Indels 88; Gaps 16;

QY 48 AQLALHRCPLAVCLAQQLPEDE--GQSSSVFED----WKLVDKSGW-----LA 95
DB 84 SSLYNLKFAHAPALPEPAPAPPEGSPVHGSPKDSFGELSRATIRLLELDQERCFLLS 143
QY 96 SVRALQQLQWHDHPRGVGRGTGL-----VERSELAFLKSPGDILTA 139
DB 144 EIEEKEKMYYSQLOQLSRDELPHVDTFSQMMDLIRQOLEPE--AQHIRS-----LM 197

QY 140 EKKDQICDFLYGRVQAREQALRLRTFOAFHSAVPSGCPCEQODEERSTRLKDLIG 199
DB 198 EERGTDDEWYQAKIR-----ASRLQIDKEL-----LEAQDRVQCTEPQALLA 242
QY 200 ---RYPEEBEGQEPGEMDAQGEPPGQARLQ-----DWEQVQCDIRQFLSL 243
DB 243 VKPVAVEEGQEAEPVTHPEDEGT-PQPGNSKVEVFWLLSMLATBDQDTATTLA--MSS 299
QY 244 RPEKPEFSR-----AVARITHG-----IGSPCTPAQVYGDDRPFWRKYLHSFHALV-- 290
DB 300 SPESCVMMRRSGCLPLLQILHTEAGSVGRAGIFGARGADARM--RANVALHNTVFS 356

QY 291 ----GLATEEL 297
DB 357 QPDGGLARKEM 367

RESULT 9
JC2549
apolipoprotein F precursor - human
C:Species: Homo sapiens (man)
C>Date: 26-Jul-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004
C:Accession: JC2549; PC2384
R:Day, J.R.; Alberts, J.J.; Gilbert, T.L.; Whitmore, T.B.; McConathy, W.J.; Wolfbauer, G.
Biochem. Biophys. Res. Commun. 203, 1146-1151, 1994
A:Title: Purification and molecular cloning of human apolipoprotein F.
A:Reference number: JC2549; MUID:94380022; PMID:8093033
A:Accession: JC2549
A:Molecule type: mRNA
A:Residues: 1-308 <DAY>
A:Cross-references: UNIPROT:Q13790; UNIPARC:UPI0000125C59; GB:L27050
A:Accession: PC2384
A:Molecule type: protein
A:Residues: 147-176;233-247 <DA2>
A:Cross-references: UNIPARC:UPI0000179D12; UNIPARC:UPI0000179D13
C:Genetics:
A:Gene: GDB:APOF
A:Cross-references: GDB:391034; OMIM:107760
A:Map position: 12pter-12qter
C:Keywords: glycoprotein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-146/Domain: propeptide #status predicted <PRO>
F:147-303/Product: apolipoprotein F #status predicted <MAT>
F:164-184/Domain: hydrophobic #status predicted <HY2>
F:207-227/Domain: hydrophobic #status predicted <HY3>
F:274-294/Domain: hydrophobic #status predicted <HY4>
F:100-121,249/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match 5.5%; Score 88.5; DB 2; Length 308;
Best Local Similarity 23.7%; Pred. No. 53;
Matches 71; Conservative 34; Mismatches 103; Indels 91; Gaps 16;

QY 14 IETLLCYLEIHP-----HHWELIATYTHCHLNCPG--GPAQLQALHRCPC 58
DB 4 VELLICVLLHPVATSVGKQTNVLMHPPLSLSESGTSSDPLSCQFLHPSLPEFSMAP 63

QY 59 ----PLAVCLAQQLPEDEPGSSSVFEDMYKLVDKSGWELASVRALCOLQWHDHPRGV 114
DB 64 LPEKVLALRNALAE-----AGCCADV-----MAL-----QLQL---YRQGG 98

QY 115 RRGVGLVVESELAFLHRSFGDILAEKQDQICDFLYGRVQ--AREQALRLRRTF----- 168
DB 99 VNAIVQVLIQ-----HLRGIQKGRSTERNVSEVALSALQILAREQDSTGRVGRSLPTED 152

QY 169 ----QAFHSA--FPSCGP-----CLEQDEERSTRLKDLGRYPFEE 206
DB 153 CENEKQAVHNWVQLRPGVGFYVNLGTALYYATQNCIGKARERRDQADILDGYLLMTMA 212

QY 207 GQE--PGEMDAQEPG-----QARLQWEDQVRCDIRQFLSLRPE-EKFSRAVARI 257
DB 213 GMSGGFWGLAISALKEALRSVQQLIYVYQDDKDAVISQ-----PETTEGRLAISDV 266

RESULT 10
E71282
probable NH(3)-dependent NAD(+) synthetase (nads) - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: E71282
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; McDo
rthy, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: E71282
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-679 <COL>
A:Cross-references: UNIPROT:O83759; UNIPARC:UPI000012PC99; GB:AE001249; GB:AE000520; NID:
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0780

Query Match 5.4%; Score 87; DB 2; Length 679;
Best Local Similarity 24.6%; Pred. No. 18;
Matches 44; Conservative 10; Mismatches 67; Indels 58; Gaps 7;

QY 31 LLAATTYHC-----RLNCPGCGPAQLQALHRCPL--AVCLAQQLPEDEPGG----- 75
DB 297 IFPAASKAHATLRBRPVCPPFPAPFOKSDAVPPLTGAVCLAVSAPSDTDGFLQRTIDL 356

QY 76 ----SSVFERDWKLVDSMGWELASVRALCOLQWHDHPRGV 114
DB 357 AAQGVALLRLEHMGCRLLVGVSGVDSACALLICARALDPLSTART--QLYALTLPGFGT 414

QY 115 RRGT-GVLVEFSR-----LAFHLRSPGDLTAEKQDQICDFLYGRVQARE 158
DB 415 TSGTKGAQGEFARALGCTVQGISASAVTHNLHDIGHTM-----QQCDGYENNAQARER 468

RESULT 11
C75499
cytidine/deoxycytidylate deaminase/nudix/methyltransferase domains protein - Deinococcus
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: C75499
R:White, O.; Bisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; f
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: C75499
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-548 <WHI>
A:Cross-references: UNIPROT:Q9BRM3; UNIPARC:UPI000003D09; GB:AE001918; GB:AE000513; NID:

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0603

A:Map position: 1

Query Match 5.3%; Score 85.5; DB 2; Length 548;

Best Local Similarity 22.1%; Pred. No. 19;

Matches 73; Conservative 41; Mismatches 135; Indels 81; Gaps 16;

8 DMEPAITLTLCLYLPHHMLE-LIATTTTCRLNCGGPAQ--LQALHRCP-PLAVG 63

61 DLAAHAIHALSVPLRRPESLSTWVLTTPVPCP-QCGAVAVMSGIRGSAAPDPWGC 119

64 LAQQLPDPGSSSVFPMVTLVDSMGWELA-----SVRALCOLMDHEPTGVRRGTG 119

120 -ARLITDDP-----YSSKGRVSRAPBPLQRAALRLMLVALLBGRPREDR 165

120 VLVEFSBLAFHLSPGDL-----TAEKDQICDPLVGRVQARERQALRLRT 167

166 LLQFSRYKADIKARELHAGTLARLRSGAGLDELTELLGALPLEMDLVLELS-- 223

168 FOAFHSVAPSCGPCIQODERSTRKDLGR---YFEEERG-----QEPGG 212

224 -PARRTAFAPDLSPGLER-----TGRACAMIEREDGFVLTARATGWTLPGGG 270

213 MEDDAGPREGQARLDWED-QVRCIDRQPLSLRPEKFSRAVARIFHGIGSPCYPAQY 271

271 IEPGETPEGAAYR-EGMEEVGARCEV-----AGAGWTLDDGSSGVCVPLRVL 316

272 G-QDRRFMRKYHLSPHALVGLATEELLQV 300

317 TLESSEPRGPLVWNPALPMADVDQLROY 346

RESULT 12

T00359

hypothetical protein KIA0680 - human

C:Species: Homo sapiens (man)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C:Accession: T00359

R:ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.

DNA Res. 5, 169-176, 1998

A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete

A:Reference number: Z14142; MUID:98403880; PMID:9734811

A:Accession: T00359

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-634 <ISH>

A:Cross-references: UNIPROT:O75167; UNIPARC:UPI000006060AA; EMBL:AB014580; NID:93327173;

A:Experimental source: brain; clone HK02746

C:Genetics:

A:Note: KIA0680

Query Match 5.3%; Score 85.5; DB 2; Length 634;

Best Local Similarity 22.7%; Pred. No. 23;

Matches 51; Conservative 26; Mismatches 91; Indels 57; Gaps 9;

58 PPLAVLCAQLPDPGQSSSVFDMVKLVDSMGWELASVRALCOLQMDHEP-RTGVR 115

327 PPAVAPASPLAPLPLEDQITASDIPVVLVSAGADLPVSLDPSQLMAEBETRTTLY 386

116 RGVGVV-----PSELAFHLSRPGDLTAERKQICDPLVGRVQARERQALR 163

387 SGTGLVNRBNACFTTKELIGTVPOLTPG-LMGSSSSP-----SASEDGG--- 434

164 LRTTFQAFHSVAPSCGPCIQODERSTRKDLGRVFEERGQEPGEMD----- 215

435 -HREYQANS---DSGPIILYTDDEBD-----EDDGGSGSALASIRRRDYL 479

216 --AAGEPFGARLD-----WEDQVRCIDRQPLSLRPEKFSR 252

480 AITLGNRPSCKELEDNLIQRTSEERQELRQOIGTKLVRLSOR 524

RESULT 13

B95878

probable adenylate cyclase (EC 4.6.1.1) [imported] - Sinorhizobium meliloti (strain 1021)

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C:Accession: B95878

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmeier, J.; Chain, P.; Votholler, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: B95878

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-637 <KDS>

A:Cross-references: UNIPROT:Q92WN9; UNIPARC:UPI00000CB49C; GB:AL591985; PIDN:CAC48690.1;

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LeJaure,

hebaull, P.; Vandenbol, M.; Votholler, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: cyar7; SMD20300

A:Genome: plasmid

C:Keywords: phosphorus-oxygen lyase

Query Match 5.3%; Score 85; DB 2; Length 637;

Best Local Similarity 21.9%; Pred. No. 25;

Matches 52; Conservative 30; Mismatches 77; Indels 78; Gaps 10;

51 QALAHRCPLAVLCAQLPDPGQSSSVFDMVKLVDSMGWELASVRALCOLQMD-HE 109

397 RALLHRA-----LELDPGVAARANLCMTIYVFPV-QNLSGRATTTDVGTSJE 444

110 PRTGVRGTGVIVSELAFLHRSFGDLTAERKQICDPLVGRVQARER----- 158

445 AAOAVRLDPLNLAAQFQVLSFGLSATGDYP-----GAMQAAQRAVELNPDPDS 492

159 -QALARLRTFOAFHSVAPSCGPCIQODERSTRKDLGRVFEERGQ---EPGWE 214

493 LMAALAAQVRFESYDAV-----QNAERARLRHPMAEYITTYTGQALYAAGRUD 542

215 DAQGEPEGARLDWEDQV--RCDIR-----QFLSLRPEKFSRAVARI 257

543 EA-----DEVLRCLIRAPQEAQCLIRTVASQRGVDEGAQRTMARL 585

RESULT 14

UC7753

ring finger B-box coiled-coil protein, GOA - human

C:Species: Homo sapiens (man)

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: UC7753

R:Vandepuute, D.A.A.; Melje, C.B.; van Darstel, M.; Ileenstra, S.; IJlst-Keizers, H.; Das,

Biochem. Biophys. Res. Commun. 286, 574-579, 2001

A:Title: GOA, a novel gene encoding a ring finger B-box coiled-coil protein, is overexpres

A:Reference number: UC7753; MUID:21402356; PMID:11511098

A:Accession: UC7753

A:Molecule type: DNA

A:Residues: 1-638 <VAN>

A:Cross-references: UNIPROT:Q96LD4; UNIPARC:UPI000013CE62; GB:AY026763

C:Comment: This protein, a member of the subfamily of the ring finger proteins, has an in

as well.

C:Genetics:

A:Gene: goa

A:Map position: 17q24-25

Query Match 5.3%; Score 85; DB 2; Length 638;

Best Local Similarity 22.6%; Pred. No. 25;
Matches 74; Conservative 29; Mismatches 104; Indels 120; Gaps 18;

Search completed: December 27, 2005, 21:54:08
Job time : 14.4 secs

QY 47 PAQQLAHRCPPPLAVCLAAQQLPEDPGQSSSVFEDMYKL-----VDSM 90
Db 183 PRHRLPERYCRARVCLCEACAAQEHGHELVLEQERALQEAQSKVLSAVEDRMDEL 242
QY 91 GWELASVRRALCOLQMDH-BPRGTGVR-----GTGYL--VEFSELAFLHRSRG 135
Db 243 GAGIAQSRRTVALIKSAVAERERYRLPADAAAALQGFOTVLGFTIEGEMAMIGRSQG 302
QY 136 DLTAEEKDQICDPLYGVRQAREQALARTRTFQAF--HSAVF-----PSCG 180
Db 303 DL-----RRQEQRSLSRARQNLGVPREADSVFLQELALRLALEDCG 348
QY 181 PCLSEQDBERSTR-----LKDLIG---RYFEEBQEPGQMEDAQPE---PGQAR 225
Db 349 PGPGPPRELSPFTKSSQAVRAVRDMLAVACVQWEQLRG--PGNED--GPQKLDEADAE 404
QY 226 LQDWE-----DQYRCDIROFLSLRPEKFSRAVARIFHGIG 262
Db 405 PQDLESTNLESEAPRDYFLKPAITVDLDSPTAD--KFLQL-----FGTKGVKRVLCPIN 457
QY 263 SPCTPA-----QYVQ--DR--RFW 278
Db 458 YPLSPTRFTHCQVGLGEGALDRGTYW 484

RESULT 15

T42692
hypothetical protein DKFZp434D0428.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42692
R:Bioecker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, U.; Wiemann, S.
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z22230
A:Accession: T42692
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-476 <AAA>
A:Cross-references: UNIPROT:Q9UFA3; UNIPARC:UPI0000070625; EMBL:AL133088
A:Experimental source: adult testis; clone DKFz434D0428
C:Genetics:
A>Note: DKFz434D0428.1

Query Match 5.2%; Score 84.5; DB 2; Length 476;
Best Local Similarity 22.2%; Pred. No. 19;
Matches 69; Conservative 26; Mismatches 117; Indels 99; Gaps 13;

QY 3 TVQALDMPERAIETLLCYLEH-----PHW--LELLATYTHCRINCPGGRQL 50
Db 5 TVDTLLASKSSCAGILCRTLAHLELOPLQRPSPWPQASILGATVTVLRLL-CDGSAAPA 63
QY 51 QAL-AHRCPPPLAVCLAAQQLPEDPGQSSSVFEDMYKLVDMSGWELASVRRALCOLQMDHE 109
Db 64 SSVGGHLCGTLAGCVRVQ-----RAALDF-----LGTLSQGTG 96
QY 110 PRITGVRGTGLVFESELAFLHRSRGDLTAEEKDQICDPLYGVRQAREQALARTRTFQ 169
Db 97 PQELVTQALAVLLELC-----LESPPGSSPTVLKKAFOATLAWLSSPTKPCSDLGPIIP 150
QY 170 AFHSAVP-----SCGFLSEQDBERSTRLLKDLIGRYFEEBQEPGQMED-----AQ 217
Db 151 QFLNELFPVLQKRLCHPCWEVRDS--ALEFLTQLSRHW-----GGQADFRCALIASB 200
QY 218 GPERGQARLQDWMEDQVR-----CDIROFLSLR-----PEE 247
Db 201 VFQALQLQLQDPEBSYVRASAVTAMQQLSSQGLHAFTSBEHAERQSLFLELLHITLSVDSB 260
QY 248 KFSRAVARIRF 258
Db 261 GFPRRAVMQVF 271

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:39:51 / Search time 58.8 Seconds

(without alignments)
3623.632 Million cell updates/sec

Title: US-09-889-325-4_COPY_907_1208
Perfect score: 1617
Sequence: 1 QLTQALMPEALFTLLCY.....HLSPHALVGLATBELGLQVAR 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_05.80:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------------|--------------------|
| 1 | 1617 | 100.0 | 1208 | 1 RECD4_HUMAN | 094761 homo sapien |
| 2 | 1144.5 | 70.8 | 1216 | 1 RECD4_MOUSE | 094761 mus musculu |
| 3 | 833 | 51.5 | 1500 | 2 Q4JX8_XENLA | 04jx8 xenopus lae |
| 4 | 753.5 | 46.6 | 910 | 2 Q4RLC3_TETNG | 04rlc3 tetraodon n |
| 5 | 464.5 | 28.7 | 1058 | 2 Q7PM8_ANOGA | 07pm8 anopheles g |
| 6 | 427 | 26.4 | 1530 | 2 Q9NH11_DROME | 09nh11 drosophila |
| 7 | 427 | 26.4 | 1579 | 2 Q9VSE6_DROME | 09vse6 drosophila |
| 8 | 186 | 11.5 | 941 | 2 Q9C6N0_ARATH | 09c6n0 arabidopsis |
| 9 | 184.5 | 11.4 | 870 | 2 Q8W028_ARATH | 08w028 arabidopsis |
| 10 | 169 | 10.5 | 874 | 2 Q7FPH0_ORYSA | 07fph0 oryza sativ |
| 11 | 169 | 10.5 | 927 | 2 Q7X829_ORYSA | 07x829 oryza sativ |
| 12 | 107 | 6.6 | 755 | 1 TR1S6_HUMAN | 09bx22 homo sapien |
| 13 | 107 | 6.6 | 809 | 2 Q6ED36_HUMAN | 09ed36 homo sapien |
| 14 | 107 | 6.6 | 900 | 2 Q6NR70_HUMAN | 06nr70 homo sapien |
| 15 | 107 | 6.6 | 1029 | 2 Q6UX74_HUMAN | 06ux74 homo sapien |
| 16 | 107 | 6.6 | 1377 | 2 Q9P2A8_HUMAN | 09p2a8 homo sapien |
| 17 | 106.5 | 6.6 | 1166 | 2 Q96JDA_HUMAN | 096jda homo sapien |
| 18 | 106.5 | 6.6 | 1681 | 2 Q8NDJ4_HUMAN | 08ndj4 homo sapien |
| 19 | 106 | 6.6 | 253 | 2 Q6ZWH8_HUMAN | 06zwh8 homo sapien |
| 20 | 105 | 6.5 | 965 | 2 Q2Z5T7_THET2 | 02z5t7 thermus the |
| 21 | 102.5 | 6.3 | 447 | 2 Q753C3_ASNGO | 0753c3 ashyia goos |
| 22 | 101.5 | 6.3 | 4678 | 2 Q468H2_TETNG | 0468h2 tetraodon n |
| 23 | 101 | 6.2 | 966 | 2 Q5S1S6_THET8 | 05s1s6 thermus the |
| 24 | 98 | 6.1 | 678 | 1 K0323_HUMAN | 015037 homo sapien |
| 25 | 97.5 | 6.0 | 1592 | 2 C01583_CAEBL | 001583 caenorabdi |
| 26 | 97 | 6.0 | 914 | 2 Q6MR18_DROTK | 06mr18 drosophila |
| 27 | 96.5 | 6.0 | 976 | 1 XPR6_YARLI | 042781 yarrowia li |
| 28 | 96.5 | 6.0 | 1418 | 2 Q5R129_BRABE | 05r129 brachydanio |
| 29 | 96 | 5.9 | 484 | 2 Q81Z03_HUMAN | 081z03 homo sapien |
| 30 | 95.5 | 5.9 | 563 | 2 Q5KG20_CRYNB | 05kg20 cryptococcu |
| 31 | 95.5 | 5.9 | 655 | 2 Q9FBR7_STRCO | 09fbr7 streptomyce |

| | | | | | |
|----|------|-----|------|-----------------|--------------------|
| 32 | 95.5 | 5.9 | 1469 | 2 Q6PD04_MOUSE | 06pd04 mus musculu |
| 33 | 95.5 | 5.9 | 1517 | 2 Q6KAS1_MOUSE | 06kas1 mus musculu |
| 34 | 95 | 5.9 | 1582 | 2 O50437_MYCTU | 050437 mycobacteri |
| 35 | 95 | 5.9 | 2085 | 2 Q7U0G2_MYCBO | 07u0g2 mycobacteri |
| 36 | 95 | 5.9 | 2101 | 2 Q8VKS2_MYCTU | 08vks2 mycobacteri |
| 37 | 94.5 | 5.8 | 298 | 1 B1RC7_HUMAN | 096ca5 homo sapien |
| 38 | 94.5 | 5.8 | 498 | 2 Q7X3G4_THIRO | 07x3g4 thioscapa r |
| 39 | 94.5 | 5.8 | 619 | 2 Q4K3N4_PSEPF5 | 04k3n4 pseudomonas |
| 40 | 94 | 5.8 | 570 | 2 Q55RL5_CRYNE | 055rl5 cryptococcu |
| 41 | 94 | 5.8 | 638 | 2 Q7U8M5_SYNPX | 07u8m5 synchococc |
| 42 | 94 | 5.8 | 798 | 2 Q5YTF9_NOCFA | 05ytf9 noctaria fa |
| 43 | 93.5 | 5.8 | 356 | 2 Q7S474_NEUCR | 07s474 neurospora |
| 44 | 93.5 | 5.8 | 610 | 2 Q41XN8_AZOVI | 041xn8 azotobacter |
| 45 | 93.5 | 5.8 | 1557 | 2 Q4S161_TETNG | 04s161 tetraodon n |

ALIGNMENTS

| RESULT 1 | RECD4_HUMAN | STANDARD; | PRT; 1208 AA. |
|----------|--|-----------|---------------|
| AC | 094761; Q96DW2; Q96F55; | | |
| DT | 30-MAY-2000 (Rel. 39, Created) | | |
| DT | 30-MAY-2000 (Rel. 39, Last sequence update) | | |
| DT | 13-SEP-2005 (Rel. 48, Last annotation update) | | |
| DE | ATP-dependent DNA helicase Q4 (EC 3.6.1.-) (RecQ protein-like 4) | | |
| DE | (RecQ4) (RTS). | | |
| GN | Name=RECD4; Synonym=RECD4; | | |
| OS | Homo sapiens (Human). | | |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; | | |
| OC | Homo | | |
| OX | NCBI_TaxID=9606; | | |
| RN | [1] | | |
| RP | NUCLEOTIDE SEQUENCE [mRNA], AND TISSUE SPECIFICITY. | | |
| RC | TISSUE=Testis; | | |
| RX | MEDLINE=99097344; PubMed=9878247; DOI=10.1006/geno.1998.5595; | | |
| RA | Kitao S., Ohnogi I., Ichikawa K., Goto M., Furuchi Y., Shimamoto A.; | | |
| RT | "Cloning of two new human helicase genes of the RecQ family: biological significance of multiple species in higher eukaryotes."; | | |
| RL | Genomics 54:443-452(1998). | | |
| RN | [2] | | |
| RP | NUCLEOTIDE SEQUENCE [GENOMIC DNA], SUBCELLULAR LOCATION, AND DISEASE. | | |
| RC | TISSUE=Lymph, and Placenta; | | |
| RX | MEDLINE=22021764; PubMed=12477932; DOI=10.1073/pnas.242603899; | | |
| RA | Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G., | | |
| RT | Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., | | |
| RL | Altechul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., | | |
| RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., | | |
| RA | Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., | | |
| RA | Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., | | |
| RA | Browstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., | | |
| RA | Bash S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullaly S.J., | | |
| RA | Rosa S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., | | |
| RA | Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., | | |
| RA | Villalon D.K., Mueny D.W., Sodergren E.J., Lu X., Gibbs R.A., | | |
| RA | Faney J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A., | | |
| RA | Whitling R.W., Madan A., Young A.C., Green E.D., Dickson M.C., | | |
| RA | Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M., | | |
| RA | Butterfield J.S.N., Krzywicki M.I., Skalska U., Smalhus D.E., | | |
| RA | Scharch A., Schein J.E., Jones S.J.M., Marra M.A., | | |
| RT | "Generation and initial analysis of more than 15,000 full-length human | | |
| RL | and mouse cDNA sequences."; | | |
| | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). | | |

[4]
 RN INDUCTION
 RP PubMed=11032027; DOI=10.1038/sj.onc.1203841;
 RX Kababe T., Tsuyama N., Kiteo S., Nishikawa K., Shimamoto A.,
 RA Shiroatori M., Matsunoto T., Anno K., Sato T., Mitani Y., Seki M.,
 RA Shimoto T., Goto M., Ellis N.A., Ide T., Furutachi Y., Sugimoto M.,
 RT "Differential regulation of human RecQ family helicases in cell
 transformation and cell cycle.";
 RL Oncogene 19:4764-4772(2000).
 RN DISEASE.
 RP PubMed=12952869; DOI=10.1093/hmg/ddg306;
 RA Siitonen H.A., Kopra O., Kaeerlaeinen H., Haravuori H., Winter R.M.,
 RA Saeenenen A.-M., Peltonen L., Kesilae M.,
 RT "Molecular defect of RAPADILINO syndrome expands the phenotype
 spectrum of RECQL diseases.";
 RL Hum. Mol. Genet. 12:2837-2844(2003).
 RN [6]
 RP INTERACTIONS WITH UBR1 AND UBR2, SUBCELLULAR LOCATION, FUNCTION, AND
 RT IDENTIFICATION BY MASS SPECTROMETRY.
 RX PubMed=15317757; DOI=10.1093/hmg/ddh269;
 RA Yin U., Kwon Y.T., Varshavsky A., Wang W.,
 RT "RECQL4, mutated in the Rothmund-Thomson and RAPADILINO syndromes,
 interacts with ubiquitin ligases UBR1 and UBR2 of the N-end rule
 pathway.";
 RL Hum. Mol. Genet. 13:2421-2430(2004).
 RN [7]
 RP VARIANTS ASP-267 AND GLN-1005.
 RX PubMed=12601557; DOI=10.1007/s100830030016;
 RA Roverai G., Beghini A., Zamburo G., Paradisi M., Larizza L.,
 RT "Identification of two novel RECQL4 exonic SNPs and genomic
 characterization of the IVS12 minisatellite.";
 RL J. Hum. Genet. 48:107-109(2003).
 RN [8]
 RP VARIANTS GLN-355; SER-441 AND 857-CYS--THR-858 DEL.
 RX PubMed=15221963; DOI=10.1002/ijc.20269;
 RA Nishijo K., Nakayama T., Aoyama T., Okamoto T., Ishibe T., Yasura K.,
 RA Shima Y., Shibata K.R., Tsuboyama T., Nakamura T., Toguchida U.,
 RT "Mutation analysis of the RECQL4 gene in sporadic osteosarcomas.";
 RL Int. J. Cancer 111:367-372(2004).
 CC -1- FUNCTION: DNA-dependent ATPase. May modulate chromosome
 segregation.
 CC -1- SUBUNIT: Interacts with UBR1 and UBR2.
 CC -1- SUBCELLULAR LOCATION: Nuclei and cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
 thymus and testis.
 CC -1- INDUCTION: Up-regulated in actively dividing cells.
 CC -1- DISEASE: Defects in RECQL4 are a cause of Rothmund-Thomson
 syndrome (RTS) [MIM:268400]. A disease characterized by
 dermatological features such as atrophy, pigmentation, and
 telangiectasia and frequently accompanied by juvenile cataract,
 saddle nose, congenital bone defects, disturbances of hair growth,
 and hypogonadism.
 CC -1- DISEASE: Defects in RECQL4 are a cause of RAPADILINO syndrome
 [MIM:266280]. A disease characterized by radial and patellar
 aplasia or hypoplasia.
 CC -1- SIMILARITY: Belongs to the helicase family. RecQ subfamily.
 CC -1- DATABAS: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 WWW="http://www.inbioigen.fr/services/chromocancer/Genes/RECQL4ID285.html".
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: AB006532; BAA74453.1; -; mRNA.
 CC EMBL: AB026546; BAA68893.1; -; Genomic_DNA.
 CC EMBL: BC011602; AAH11602.2; -; mRNA.
 CC EMBL: BC013277; AAH13277.2; ALT_INIT; mRNA.
 CC HSSP: P15043; IOY.
 CC Ensemble: ENSG00000160957; Homo sapiens.
 CC DR HGNC: 9949; RECQL4.

DR MIM: 603780; -;
 DR MIM: 268400; -;
 DR MIM: 266280; -;
 DR GO: GO:0003678; P:DNA helicase activity; TAS.
 DR GO: GO:0003678; P:DNA helicase activity; TAS.
 DR GO: GO:0006281; P:DNA repair; TAS.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR002464; DEAH_box.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR004589; RecQ.
 DR PANTHER: PTHR13710; RecQ; 1.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR TIGRFAMs: TIGR00614; recQ_fam; 1.
 DR PROSITE: PS00690; DEAH_ATP_HELICASE; FALSE NEG.
 KW ATP-binding; Catalytic; Helicase; Hydrolyase; Nuclear protein;
 KW Nucleotide-binding; Polymorphism.
 NP BIND 502 509
 FT NP BIND 502 509
 FT MOTIF 605 608
 FT NP BIND 267 267
 FT VARIANT 355 355
 FT FT 441 441
 FT VARIANT 441 441
 FT FT 857 858
 FT VARIANT 1005 1005
 FT FT 1005 1005
 SQ SEQUENCE 1208 AA; 133077 MW; CB809A765AB48A1 CRC64;
 Query Match 100.0%; Score 1617; DB 1; Length 1208;
 Best Local Similarity 100.0%; Pred. No. 1.9e-125;
 Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QLTVALDMEBEAETLLCYLSEHPHMLBLATTTTCRLNCGSGPAQIALAHRCPL 60
 DB 907 QLTVALDMEBEAETLLCYLSEHPHMLBLATTTTCRLNCGSGPAQIALAHRCPL 966
 QY 61 AVCLAQQLPEDPGGSSSVFEDWVKLVDSMGWELASVRALCOLQWHPPTGVRGRTGV 120
 DB 967 AVCLAQQLPEDPGGSSSVFEDWVKLVDSMGWELASVRALCOLQWHPPTGVRGRTGV 1026
 QY 121 LVFSESLAFLHRSBGDTJAEKKQICDPLVGRVQAREQALRLRTFOAFHSAVPSCG 180
 DB 1027 LVFSESLAFLHRSBGDTJAEKKQICDPLVGRVQAREQALRLRTFOAFHSAVPSCG 1086
 QY 181 PCLEQDEERSTRKDLGRFYFEEBEGQBGWMDAGPSPGARLQDWDQVCDIRQF 240
 DB 1087 PCLEQDEERSTRKDLGRFYFEEBEGQBGWMDAGPSPGARLQDWDQVCDIRQF 1146
 QY 241 LSLRPEKFSRAVARIFHGIGSPCYPAOYVGODRRFMGRKYLHLSFALVGLATEELLQV 300
 DB 1147 LSLRPEKFSRAVARIFHGIGSPCYPAOYVGODRRFMGRKYLHLSFALVGLATEELLQV 1206
 QY 301 AR 302
 DB 1207 AR 1208
 RESULT 2
 RECO4 MOUSE STANDARD; PRT; 1216 AA.
 AC 075NR7; 076MT1; 099PV9;
 DT 13-SEP-2005 (Rel. 48, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE ATP-dependent DNA helicase Q4 (BC 3.6.1.-) (RecQ protein-like 4).
 GN Name=RecQ4; Synonyms=Recq4;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

CC Muroidea; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA / mRNA] (ISOFORM 1) AND NUCLEOTIDE
RP SEQUENCE [mRNA] OF 16-1115 (ISOFORM 2).
RX PubMed=1167012;
RA Ohnata T., Araki R., Fukumura R., Kuroiwa A., Matsuda Y., Tatsumi K.,
RA Abe M.;
RT "Cloning, genomic structure and chromosomal localization of the gene
RT encoding mouse DNA helicase Recq helicase protein-like 4.";
RL Gene 261:251-258(2000).
RN [2]
RN FUNCTION:
RP PubMed=12915449; DOI=10.1093/jmg/ddg254;
RA Hoki Y., Araki R., Fujimori A., Ohnata T., Koseki H., Fukumura R.,
RA Nakamura M., Takahashi H., Noda Y., Kito S., Abe M.;
RT "Growth retardation and skin abnormalities of the Recql4-deficient
RT mouse";
RL Hum. Mol. Genet. 12:2293-2299(2003).
RN [3]
RN DEVELOPMENTAL STAGE.
RP PubMed=12952869; DOI=10.1093/jmg/ddg306;
RA Siitonen H.A., Kopra O., Kaeareiseinen H., Haravuori H., Winter R.M.,
RA Saalenanen A.-M., Peltonen L., Kesiläe M.;
RT "Molecular defect of RAPADILINO syndrome expands the phenotype
RT spectrum of RECQL diseases.";
RL Hum. Mol. Genet. 12:2837-2844(2003).
RN [4]
RN FUNCTION:
RP PubMed=15703196; DOI=10.1093/jmg/ddi1075;
RA Mann M.B., Hodges C.A., Barnes E., Vogel J., Luo G.;
RA "Defective sister-chromatid cohesion, aneuploidy and cancer
RT predisposition in a mouse model of type II Rothmund-Thomson
RT syndrome";
RL Hum. Mol. Genet. 14:813-825(2005).
RN -1- FUNCTION: DNA-dependent ATPase. (By similarity). May play a role in
RN development of the palate and the limbs. May modulate chromosome
RN segregation.
CC -1- SUBUNIT: Interacts with UBR1 and UBR2 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q75NR7-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q75NR7-2; Sequence=VSP_015177;
CC -1- DEVELOPMENTAL STAGE: Not expressed at E12.5. Expressed at E15.5-
CC E18.5, with highest levels in chondrocytes of developing bone and
CC cartilage and immature proliferating enterocytes of intestine.
CC MISCELLANEOUS: Absence of the protein causes early embryonic
CC lethality. Transgenic mice with exon 13-deleted RECQL4 are severely
CC growth-retarded and show high (95%) perinatal lethality. They
CC exhibit various skin, bone, intestine, tooth and thymus
CC abnormalities and premature aging features, but have normal
CC sensitivity to IR and UV irradiation. In contrast, transgenic mice
CC expressing a truncated form of RECQL4 exhibit mild perinatal
CC lethality, no growth defect, but show defects of the skin and
CC skeleton, aneuploidy and increased cancer susceptibility.
CC -1- SIMILARITY: Belongs to the helicase family. Recq subfamily.
CC -1- SIMILARITY: Contains 1 CCHC-type zinc finger.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, AB039882; BADI1131.1; -, mRNA.
DR EMBL, AB175741; BADI14289.1; -, mRNA.
DR EMBL, AB045259; BAB32696.1; -, Genomic_DNA.
DR HSSP, P15043; IOY.
DR Ensemble, ENSMUSG00000033762; Mus musculus.
DR MGJ, MGJ:1931028; Recq14.

DR PROSITE; PS00690; DEAH ATP HELICASE; FALSE_NEG.
 DR PROSITE; PS00158; ZF_CCHC; 1.
 KM Alternative splicing; ATP-binding; Helicase; Hydrolase; Metal-binding;
 KW Nuclear protein; Nucleotide-binding; zinc; Zinc-finger.
 FT ZN FING 393 410 CCHC-type.
 FT NP BIND 519 526 ATP (Potential).
 FT MOTIF 627 630 DEAH box.
 FT VARSP LIC 843 843 O -> OVGSSTPPODPRGSTTPRPQPOLSLCLPVSCRP
 GPGSGSVHTVMQ (in isoform 2).
 /FTIDE_VSP_015177.
 ERIQNGWL -> RKDERVA (in Ref. 1);
 BAD14289).

Query Match 70.8%; Score 1144.5; DB 1; Length 1216;
 Best Local Similarity 73.2%; Pred. No. 3.9e-86;
 Matches 218; Conservative 26; Mismatches 53; Indels 1; Gaps 1;

| | | | |
|----|------|---|------|
| Qy | 1 | QLTVQALDMBEAEIETLLCYLELHPHMLBLATTTTHCRINCPCGPAOLQALAHRCPL | 60 |
| Db | 916 | QSTVQALDMTBEALETLLCTELHPRMDELPTTYAQCHLHCIGSAGIQALAHRCPL | 975 |
| Qy | 61 | AVCIAQQLPEDPGSSGVPEFDWKVLVDMSGMELASVRRLCOLQMDEHPRTGYRRTGV | 120 |
| Db | 976 | AACQAKMPKDTQSGRSLSLEGVELADMSMGWKLASVRAQLHLQKWDPBPKKGAAOOGTV | 1035 |
| Qy | 121 | LVSSELAFLHSRGDLTAERKDCIDPLVGRVGARERQALRLRRTFQAFHSAVPFGCG | 180 |
| Db | 1036 | LVKFSELAFLHSRGDLTDREKDCIDPLVNRVARRHKAHLHLQMSKARSAVAFPSCG | 1095 |
| Qy | 181 | PCTEGODEBRTRFKLDLGFRFEESGECPGEMDAOCGPFGQARLODMEOVCNDRQF | 240 |
| Db | 1096 | PCLDSNBESHNSQKTLVSYTFEEBE-BEESTMTDITDGPKKGQVQLQDMENQIRDVQL | 115 |
| Qy | 241 | LSLRPEEFKFSRAVARIFHGIGSDCYPAQYVGDDRRFMVKYLLHSFHALVGLATEELL | 298 |
| Db | 1155 | LSLRPERFSGRAVARIFHGIAS-CYPAYYGJLDRRFMRKYLLHDFFHALMILATEELL | 1212 |

RESULT 3
 QJUNK8 XENLA PRELIMINARY; PRT; 1500 AA.

| | | |
|----|---|----------------------------|
| ID | QJUNK8 XENLA | PRELIMINARY; PRT; 1500 AA. |
| AC | QJUNK8- | |
| DT | 13-SEP-2005 (TREMBLrel. 31, Created) | |
| DT | 13-SEP-2005 (TREMBLrel. 31, Last sequence update) | |
| DT | 13-SEP-2005 (TREMBLrel. 31, Last annotation update) | |
| DE | RECOLA-helicase-like protein. | |
| GN | Name=RFS; | |
| OS | Xenopus laevis (African clawed frog). | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; | |
| OC | Xenopodinae; Xenopus; Xenopus. | |
| RX | NCBI_Taxid=8355; | |
| OR | [1] | |
| RP | NUCLEOTIDE SEQUENCE. | |
| RA | Sangrathi M.N., Bernal J.A., Madine M., Philpott A., Lee J., | |
| RT | Dunphy W.G., Venkatarman A.R.; | |
| RT | "Initiation of DNA replication requires the RECOLA protein mutated in | |
| RU | Roehmund-Thomson syndrome."; | |
| RU | Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases. | |
| DR | EMBL; DQ059311; AAY89585.1; -, mRNA. | |
| KW | Helicase. | |
| SQ | SEQUENCE 1500 AA; 168706 MW; A9SF0A2BB2BE554D CRC64; | |

Query Match 51.5%; Score 833; DB 2; Length 1500;
 Best Local Similarity 54.6%; Pred. No. 4e-60;
 Matches 165; Conservative 53; Mismatches 78; Indels 6; Gaps 4

3 TVQALDMBEAEIETLLCYLELHPHMLBLATTTTHCRINCPCGPAOLQALAHRCPLAV 62

```
Db 1200 TVQSLDIREAIETMLCYLELNPSTMLEQLHPTLSHCIVVCSSGQQLRLARSPFPAV 1259
Qy 63 CLA-QQLPEDPQCGSSSVFEDMVKLVDSMGWELASVRRALCOLQMDHEPRTGVRRGT--- 118
Db 1260 CLARERLAGVHTHTVSSSTEFNVVELADSMGWVAVVKALNDLQRRRQDPDKGF-KGTGRS 1318
Qy 119 GVLPFSELAFLHRSFGDLTAEBEKQICDFLYGRVQAERQALRLKRTFOAFHSVAPPS 178
Db 1319 GVLPFSELAFLHRSFGDLTAEBEKQICDFLYGRVQAERQALRLKRTFOAFHSVAPPS 1378
Qy 179 CGPCLQODEBERSTRYKDLIGRYEFEESGQEPGEMDAQGPPEQARLQMDMDQVRCIR 238
Db 1379 SSACAEKMSARSAQQLKALLDYFEKRDTELEAKVEDD-BELOQIKYQEAQDQVRSVVR 1437
Qy 239 QFLSLRPEKFSRAVARIFHGIGSPCYPAQVYGDRFRWRKYLHLSFHALVGLATEEL 298
Db 1438 HFLSIHEERFSGALARIIFHGIGSPCYPAQVYGDRFRWRKYLHLSFHALVGLATEEL 1497
Qy 299 QY 300
Db 1498 RL 1499

RESULT 4
04RLC3_TERNG PRELIMINARY; PRT; 910 AA.
ID 04RLC3_TERNG PRELIMINARY; PRT; 910 AA.
AC 04RLC3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 21 SCAF15022, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00032550001;
OS Tetradon nigraviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neocentropomidae;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetradon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mucelli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segutrens B.,
RA Daelya C., Salenouat M., Levy M., Boudet N., Castellano S.,
RA Authouard V., Jubin C., Castell V., Kacinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cactolico L., Poulain J., De Bernardis V.,
RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan F., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetradon nigraviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01015022; CAG10809.1; -; Genomic_DNA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00098; zf-CCHC; 1.
```

```
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR SMART; SM00343; ZNF_C2HC; 1.
DR TIGRFAMs; TIGR00614; recQ_fam; 1.
DR PROSITE; PS0158; 2F_CCHC; 1.
KW ATP-binding; Helicase; Hydrolase.
FT NON TER
FT NON TER
FT SEQUENCE
SQ SEQUENCE 910 AA; 101865 MW; 7CEFF6939FB4564F CRC64;

Query Match 46.6%; Score 753.5; DB 2; Length 910;
Best Local Similarity 50.2%; Pred. No. 9.2e-54;
Matches 151; Conservative 56; Mismatches 85; Indels 9; Gaps 5;

Qy 1 QLTVALDMPREALITLLCYLELPHHMLBLATTYTHCRNLNCPGPAQALAHRCPL 60
Db 615 QQTVEALDITBEGVETILCTLELHPQRFVFLHPTLSVCKISCYDGRQLRRVTKICPV 674
Qy 61 AVCLIA-QQLPEDPQCGSSSVFEDMVKLVDSMGWELASVRRALCOLQMDHEPRTGVRRG-T 118
Db 675 AVVLARRRMAGERVESCDLEFVVEVADTWGMQLPLVKGRLOQWS---TGQGGGRS 730
Qy 119 GVLPFSELAFLHRSFGDLTAEBEKQICDFLYGRVQAERQALRLKRTFOAFHSVAPPS 178
Db 731 GVLPFSELAFLHRSFGDLTAEBEKQICDFLYGRVQAERQALRLKRTFOAFHSVAPPS 790
Qy 179 CGPCLQODEBERSTRYKDLIGRYEFEESGQEPGEMDAQGPPEQARLQMDMDQVRCIR 237
Db 791 VSSQDQDLDTDRSLQKLSLSLEYFDKRRBGDRVAPEDYGLD--RKVLDMEQIRADI 848
Qy 238 RQFLSLRPEKFSRAVARIFHGIGSPCYPAQVYGDRFRWRKYLHLSFHALVGLATEEL 297
Db 849 RSFLANRSDEKFSRAVARILHIGSPCYPAQVYGDRFRWRKYLHLSFHALVGLATEEL 908
Qy 298 L 298
Db 909 I 909
```

```
RESULT 5
07PNM8_ANOGA PRELIMINARY; PRT; 1058 AA.
ID 07PNM8_ANOGA PRELIMINARY; PRT; 1058 AA.
AC 07PNM8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000004546 (Fragment).
GN ORFNames=BNSANG50000003572;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008960; EAA11913.3; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
```


RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Deicher A., Deng Z., Kays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Morklov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheller F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP NCBIOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Ceiniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Paleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT melanogaster euchromatic genome sequence."
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP NCBIOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminier J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Ceiniker S.E.,
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomic perspective."
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 RN [4]
 RP NCBIOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminier J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Ceiniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review."
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [5]
 RP NCBIOTIDE SEQUENCE.
 RG Berkeley *Drosophila* Genome Project;
 RA Ceiniker S., Carlson J., Wan K., Pfeiffer B., Friese E., George R.,
 RA Hoskins R., Stapleton M., Paleb J., Park S., Svirskas R., Smith E.,
 RA Yu C., Rubin G.,
 RT "Drosophila melanogaster release 4 sequence."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP NCBIOTIDE SEQUENCE.

RG FlyBase;
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003556; AAF50477.1; -, genomic DNA.
 DR HSSP; P15043; 10Y.
 DR Ensemble; CG7487; *Drosophila melanogaster*.
 DR FlyBase; FBGN0040290; CG7487.
 DR FlyBase; FBGN0040290; RecQ4.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0006259; P:DNA metabolism; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004589; RecQ.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00098; zf_CCHC; 1.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR TIGRFAMs; TIGR00614; recQ_fam; 1.
 DR PROSITE; PS50158; ZF_CCHC; 1.
 SQ SEQUENCE 1579 AA; 175909 MW; EA24B9B49C8FCF7D CRC64;
 Query Match 26.4%; Score 427; DB 2; Length 1579;
 Best Local Similarity 31.1%; Pred. No. 2.5e-26;
 Matches 95; Conservative 62; Mismatches 122; Indels 26; Gaps 6;
 QY 3 TVQALDMPBEAIEITLTCYELHPHMLATTTTHRLNCPSGPALQALAHRCPLAV 62
 DB 1290 TVEMLDIPAEINISITLCYMEIDPRWCISVLSAYVAKVISYGGPKYLKNAAXCEPLAM 1349
 QY 63 CLAQLLEDP- GQSSSVFPMVVLVDVSMGELLASVRAICQLQMDHPEPTGVARGVYL 121
 DB 1350 AIAQLQDKTKFKESNSITFSVDIDAIGTGNSSGVYKQLDLEW--VKVNGYKRSPTT 1407
 QY 122 VEPSELAFLHRLSPGDLTAEEKDQICDPLVGRVQAREQALRLRRTQAFHSVAFSPCGP 181
 DB 1408 VSFYDLGFRIKVPDPDFESEIDNALDPLTYRSVQERTQLQLQYVHAGLAVAVSSCGQ 1467
 QY 182 CLEEQ-DEBSSTRKLKILGRFEEBEGQ-----EPGAMEDAQGPBGQARLQDWEQVR 234
 DB 1468 CCNADPFODRGEQKALVRYVYFANDYPQDLELEIEPSNVDP-----ENII 1512
 QY 235 CDIRQPSLTPBEKFSRAVARIFPHGISPCYPAQVYGQDRRFKYLHLSFHALVGLAT 294
 DB 1513 DDVALINMTPDNTFTGRNIRARIFPHGIMSPYPAVIMGR-CKEWRARHVKVDFNRIHLHAN 1571
 QY 295 BELQ 299
 DB 1572 MATIK 1576
 RESULT 8
 ID Q9C6N0 ARATH PRELIMINARY; PRT; 941 AA.
 AC Q9C6N0.
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
 DE ATP-dependent DNA helicase RecQ, putative.
 GN Name=F28L5.4;
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 CC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsia.
 CX NCBI_Taxid=3702;
 RN [1]
 RP NCBIOTIDE SEQUENCE.
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,

RA Maiti R., Romning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Bannstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: EIF4A is both a subunit of a high molecular weight
CC protein complex involved in cap recognition and is required as a
CC single polypeptide chain for mRNA binding to ribosome. It is an
CC ATP-dependent single stranded DNA-binding protein with a sequence-
CC independent unwinding activity (helicase) (By similarity).
CC -I- SUBUNIT: eIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least EIF4A, EIF4B and EIF4G (By similarity).
CC EMBL; AC079280; AAC50580.1; -; Genomic_DNA.
DR HSSP; P15043; 10YY.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002464; DEAD box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR TIGRFAMs; TIGR00614; recQ_fam; 1.
DR PROSITE; PS00690; DEAH ATP HELICASE; UNKNOWN_1.
DR ATP-binding; Helicase; Hydrolyase; Initiation factor; Nuclear protein;
KW Protein biosynthesis.
SQ SEQUENCE 941 AA; 104156 MW; 30BC454C795522E1 CRC64;

Query Match 11.5%; Score 186; DB 2; Length 941;
Best Local Similarity 24.4%; Pred. No. 1.5e-06;
Matches 77; Conservative 45; Mismatches 126; Indels 66; Gaps 12;

QY 5 QALDMPBEAIEITLLCYLELPHHMLLATTYTHCRINCPCGPAQLALAHRCPPLA 61
DB 659 QKFDKKEEVMQITLTHLEGEVQLRMLPQINICCTLN-----FHKTSDAQNIV 707
QY 62 VCLAQQLPEDPGCGSSSVFPMVKLVDSMG---WEIASYRALCOLQWHEPRVGRKCT 118
DB 708 MLYVSSPPTLARSATVAAIILKSHVKGQGLHVFDPVAVASSICVAT-----T 755
QY 119 GVLVSESLAFLHRSFGDLTAEEKD-----QICDFLYGRVQARERQALRLRR 166
DB 756 DVLAIEIQL-----KGEVYTELKDSAFCTYTIKSKKICSLSHLTMLTEIESCKYAK 809
QY 167 -TFQAFHSVAFSPCGPCLQODEERSTRKDLGRYFEEBEGQEPGMEDAQEPFGQAR 225
DB 810 LDIMSAAVAIAISVNTSELSSGAKQTR--SLQSRIFDYFNG-----DEKDSPSKA- 859
QY 226 LODMEDQ---VRCIDRQPLSLRPEEKSSRAVAIFHGIGSPCYPAQYGGDRFRWKRY 281
DB 860 -----TONCAFIRADIKVFLQSNRQAKFTPRAIARIMHGVSPPAFNSVWSK-THFWGRY 913
QY 282 LHLSPHALVGLATEEL 297
DB 914 MNVDPRVIMEAQTEL 929

RESULT 9

Q8M028_ARATH PRELIMINARY; PRT; 870 AA.
AC Q8M028;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Helicase.

GN Name=rf1-5;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OK NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hartung F., Plachova H., Puchta H.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: EIF4A is both a subunit of a high molecular weight
CC protein complex involved in cap recognition and is required as a
CC single polypeptide chain for mRNA binding to ribosome. It is an
CC ATP-dependent single stranded DNA-binding protein with a sequence-
CC independent unwinding activity (helicase) (By similarity).
CC -I- SUBUNIT: eIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least EIF4A, EIF4B and EIF4G (By similarity).
CC EMBL; AJ421618; CAD13472.1; -; mRNA.
DR HSSP; P15043; 10YY.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003743; F:translation initiation factor activity; IEA.
DR GO; GO:0006259; P:DNA metabolism; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002464; DEAD box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004589; RecQ.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR TIGRFAMs; TIGR00614; recQ_fam; 1.
DR PROSITE; PS00690; DEAH ATP HELICASE; UNKNOWN_1.
DR ATP-binding; Helicase; Hydrolyase; Initiation factor; Nuclear protein;
KW Protein biosynthesis.
SQ SEQUENCE 870 AA; 96383 MW; 8B5ADA4F47A81723 CRC64;

Query Match 11.4%; Score 184.5; DB 2; Length 870;
Best Local Similarity 24.3%; Pred. No. 1.9e-06;
Matches 76; Conservative 45; Mismatches 119; Indels 73; Gaps 12;

QY 5 QALDMPBEAIEITLLCYLELPHHMLLATTYTHCRINCPCGPAQLALAHRCPPLA 64
DB 599 QKFDKKEEVMQITLTHLEGEVQLRMLPQINICCTLN-----FHKSSPNTLAA 647
QY 65 AQQLPEDPGCGSSSVFPMVKLVDSMG---WEIASYRALCOLQWHEPRVGRKCTVL 121
DB 648 R-----SAIYAAIILKSHVKGQGLHVFDPVAVASSICVAT-----TDVL 685
QY 122 VEFSESLAFLHRSFGDLTAEEKD-----QICDFLYGRVQARERQALRLRR-TF 168
DB 686 AEIQ-----LKKKGEVYTELKDSAFCTYTIKSKKICSLSHLTMLTEIESCKRKLDI 741
QY 169 QAFHSVAFSPCGPCLQODEERSTRKDLGRYFEEBEGQEPGMEDAQEPFGQARLQD 228
DB 742 MSSAAVAIAISVNTSELSSGAKQTR--SLQSRIFDYFNG-----DEKDSPSKA- 788
QY 229 WEDQ---VRCIDRQPLSLRPEEKSSRAVAIFHGIGSPCYPAQYGGDRFRWKRYLHL 284
DB 789 --TONCAFIRADIKVFLQSNRQAKFTPRAIARIMHGVSPPAFNSVWSK-THFWGRY 845
QY 285 SFHALVGLATEEL 297
DB 846 DFRVIMEAQTEL 858

RESULT 10

Q7FAH0. ORYSA
ID Q7FAH0. ORYSA PRELIMINARY; PRT; 874 AA.
AC Q7FAH0;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE OJ000223 09.17 protein.
Name-OJ000223 09.17;
OS Oryza sativa [japonica cultivar-group].
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
[1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
"Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
CC -!- FUNCTION: EIF4A is both a subunit of a high molecular weight
CC protein complex involved in cap recognition and is required as a
CC single polypeptide chain for mRNA binding to ribosome. It is an
CC ATP-dependent single stranded DNA-binding protein with a sequence-
CC independent unwinding activity (helicase) (By similarity).
CC -!- SUBUNIT: eIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least EIF4A, EIF4B and EIF4G (By similarity).
CC EMBL; AL606998; CAD1405.2; -; Genomic_DNA.
DR Gramene: O7FAH0; -;
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0003743; F:translation initiation factor activity; IEA.
DR GO: GO:0006259; P:DNA metabolism; IEA.
DR GO: GO:0006412; P:protein biosynthesis; IEA.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR004589; RecQ.
DR Pfam: PF00270; DEAD; 1.
DR SMART; SM00487; DEXDC; 1.
DR TIGRFAMs; TIGR00614; recQ_fam; 1.
KW Helicase; Initiation factor; Protein biosynthesis.
SQ SEQUENCE 874 AA; 96275 MW; 51B64C48617774F CRC64;
Query Match 10.5%; Score 169; DB 2; Length 874;
Best Local Similarity 20.9%; Pred. No. 3.6e-05;
Matches 68; Conservative 58; Mismatches 118; Indels 82; Gaps 12;
QY 1 QLTVALDMPBEALETLLCYLELPHNHMLELATYTHRCRLNCPGPAQIALAHRCPL 60
DB 598 ELTSRKPDKEEVLLITLTQLEIDQYIRLLPFSVYCTL-----YHNKTSPPQ 646
QY 61 AVC-----LAQOLPRDPQGSSTVSFEDMVKLVDSNGWELASRYRLCOLQMDHERTGVRR 116
DB 647 LLDKDKLILNSVLRSEMKQGHVFDIPRLANDKITMNEV-----FDHLNR----- 693
QY 117 GTGLVFRS-ELAFHLRSPG-----DLTAEEKQDIDPLYGRVQAEQALRLR 165
DB 694 -----LKFSGEISBELDPAVCYVILMRPDDFNLSAN-----LTKMLSEVSSKISKLD 743
QY 166 RTQAFHSVAPFSC-----GPCLQEQDEERSTRKLDLGRYFEEBEGQEPGME 214
DB 744 AMF-ALANFAVKGCKRTGSGSGQHTPCIQKKIME-----YFSKDDGTSEMDC- 790
QY 215 DAGPREGQARLQMDQVQKCDIRQPLSLRBEKFSRAVARIRHIGSGPCTYPAQVYGOD 274

DB 791 -----RTQLOKSSPEFLADIKVFIQSNFATFTFRAVARIMHGISSPAFBSVTWSKN 842
QY 275 RRFWRKYHLSPHALVGLATEELLQV 300
DB 843 -HFMGRVVEVDPLVWMAAKAEVLKL 867
RESULT 11
Q7X829. ORYSA
ID Q7X829. ORYSA PRELIMINARY; PRT; 927 AA.
AC Q7X829;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE OJ991113 30.2 protein.
GN Name-OJ991113 30.2;
OS Oryza sativa [japonica cultivar-group].
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
[1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Mu W., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
"Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
CC -!- FUNCTION: EIF4A is both a subunit of a high molecular weight
CC protein complex involved in cap recognition and is required as a
CC single polypeptide chain for mRNA binding to ribosome. It is an
CC ATP-dependent single stranded DNA-binding protein with a sequence-
CC independent unwinding activity (helicase) (By similarity).
CC -!- SUBUNIT: eIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least EIF4A, EIF4B and EIF4G (By similarity).
CC EMBL; AL662946; CAD14320.2; -; Genomic_DNA.
DR HSPB; P15043; 10YX.
DR Gramene: Q7X829; -;
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003677; F:DNA binding; IEA.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR GO: GO:0003743; F:translation initiation factor activity; IEA.
DR GO: GO:0006259; P:DNA metabolism; IEA.
DR GO: GO:0006412; P:protein biosynthesis; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR004589; RecQ.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR TIGRFAMs; TIGR00614; recQ_fam; 1.
KW ATP-binding; Helicase; Hydrolase; Initiation factor;
KW protein biosynthesis.
SQ SEQUENCE 927 AA; 102031 MW; 48B7BC5504C04FPA CRC64;
Query Match 10.5%; Score 169; DB 2; Length 927;
Best Local Similarity 20.9%; Pred. No. 3.9e-05;
Matches 68; Conservative 58; Mismatches 118; Indels 82; Gaps 12;
QY 1 QLTVALDMPBEALETLLCYLELPHNHMLELATYTHRCRLNCPGPAQIALAHRCPL 60

```

Db      651 ELTSKPKDILKEVLTITLTQLEIGQQYIRLLPQPSVYCTL-----YFHKTSPO 699
Oy      61 AVC-----LAQQLPEPQGGSSSVEDPWKYDMSKMEELASVRALCQLOMHEPPTGRR 116
Db      700 LLAQKDLIRLSTLNSEKMDGHYVDPRIADLKITNNEV-----FHLHK----- 746
Oy      117 GGVGVVRS-ELAFHLRSPG-----DLTAEEKQICDPLGVAOERQALRLR 165
Db      747 -----LKSSEISFEKLPAYCYVLMKPPDNLASN-----LTKMSSEVSSKISLSD 796
Oy      166 RTFOAFHSVAPFSC-----GPCLEQDEERSITRLKDLGRYEEERGQEPGME 214
Db      797 AMF-ALANFAYVKGCRKGCGSGSGSHPTCIQKKIME-----YFSKDGTSENDC- 843
Oy      215 DAQPEPQALQDMEDVRCIRPFLSRPEKSSRAVAFIFGIGSPCEPAQVYQOD 274
Db      844 -----RTQLQKSPPLQADIKVFIOGNSPFAKFTPRAVARIMGIGSSPAPSVTWSKN 895
Oy      275 RRFMRKYLHSPHALVGLATEBLQV 300
Db      896 -HFGWGVVEVDPPLVMEAKKALVPL 920

RESULT 12
TRIS6_HUMAN STANDARD; PRT: 755 AA.
AC      Q9BRZ2; Q9BRV7; Q9NZH8; Q9NAC0; Q9H031;
DT      01-FEB-2005 (Rel. 46, Created)
DR      10-FEB-2005 (Rel. 46, Last sequence update)
DT      01-MAY-2005 (Rel. 47, Last annotation update)
DE      Tripartite motif protein 56 (RING finger protein 109).
GN      Name=TRIM56; Synonyms=RNFP109;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC      Homo.
NC      NCBI_TaxID=9606;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX      MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA      Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA      Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA      Wyllie K., Sekhon M., Becker M.C., O'Leughlin M.D., Schaller M.E.,
RA      Powell G.A., Delenauy K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA      Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA      Vandrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA      Ozeraky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA      Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA      Kozlowski-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA      Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strowmatt C.,
RA      Wendt M.C., Yang S.-P., Schultz B.R., Wallis J.W., Splich J.P.,
RA      Bieri T.A., Nelson J.O., Berkowicz N., Wohlmann P.E., Cook L.L.,
RA      Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
RA      Clifton S.W., Chisoe S.L., Marra M.A., Raymond C., Haugen E.,
RA      Gillett W., Zhou Y., James R., Phelps K., Iadonco S., Buty K.,
RA      Simme E., Levy R., Clendinning J., Kaul R., Kent W.J., Furey T.S.,
RA      Baerbach R.A., Brent M.R., Keibler E., Flitcey P., Bork P., Suyama M.,
RA      Bailey J.A., Portnoy M.E., Torrente D., Chinwalla A.T., Gish W.R.,
RA      Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA      Waterston R.H., Wilson R.K.,
RT      "The DNA sequence of human chromosome 7."
RL      Nature 424:157-164(2003).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC      TISSUE=Lymph, and Placenta;
RX      MEDLINE=22382857; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Pingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner C.M., Shenmen G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Heath P.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Heien F.,
RA      Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```

```

RA      Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Pirange C.,
RA      Baka S.S., Lequellan N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA      Roark S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Botterigues Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA      Schnerker A., Schein J.E., Jones S.J.M., Marra M.A.,
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences."
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [3]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 1-269 AND 429-755.
RC      TISSUE=Spleen, and Thymoid;
RX      PubMed=14702039; DOI=10.1038/ng1285;
RA      Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA      Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA      Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA      Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA      Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA      Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA      Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA      Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA      Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA      Fujimori K., Tanai H., Kimeta M., Watanabe M., Hiraoa S., Chiba Y.,
RA      Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Houchi T.,
RA      Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA      Nomura Y., Togashi S., Komai F., Hara R., Takeuchi K., Arita M.,
RA      Imose N., Mueshino K., Yuki F., Oshima A., Sasaki N., Aotaka S.,
RA      Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA      Moriya S., Momiyama H., Satoh N., Takami S., Terasushima Y., Suzuki O.,
RA      Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA      Hishigaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA      Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA      Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA      Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,
RA      Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA      Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA      Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA      Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA      Mizushima-Sugano J., Sato H., Shirai Y., Takahashi Y., Nakagawa K.,
RA      Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA      Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT      "Complete sequencing and characterization of 21,243 full-length human
RT      cDNAs."
RL      Nat. Genet. 36:40-45(2004).
RN      [4]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 572-755.
RC      TISSUE=Lymph node;
RX      The German cDNA consortium;
RG      Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RL      [1]
CC      -1- SIMILARITY: Belongs to the TRIM/RBCC family.
CC      -1- SIMILARITY: Contains 2 B box-type zinc fingers.
CC      -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
DR      EMBL; BC005847; AAH05847.2; ALT_INIT; mRNA.
DR      EMBL; BC048194; AAH48194.1; -; mRNA.
DR      EMBL; AK092927; BAC04004.1; -; mRNA.
DR      EMBL; AK075255; BAC11500.1; ALT_INIT; mRNA.
DR      EMBL; AL512757; CAC21676.1; -; mRNA.
DR      HSSP; P29590; 1BOK.
DR      Ensemble; ENSG00000169871; Homo sapiens.
DR      HGNC; HGNC:19028; TRIM56.

```

```
DR InterPro; IPR000315; ZnF Bbox.
DR InterPro; IPR001841; ZnF_RING.
DR Pfam; PF006643; zf-B_box.2.
DR Pfam; PF000097; zf-C3HC4; 1.
DR PRINTS; PRO1406; BBOXZNFINGER.
DR SMART; SM00336; BBOX; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00119; ZF_BOX; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00889; ZF_RING_2; 1.
KM Coiled coil; Metal-binding; Repeat; Zinc; Zinc-finger.
FT ZN_FING 21 60 RING-type.
FT ZN_FING 98 149 B box-type 1.
FT ZN_FING 164 205 B box-type 2.
FT COILED 216 314 Potential.
FT COMPBIA 302 309 Poly-Ala.
FT COMPBIA 459 464 Poly-Lys.
FT CONFLICT 230 230 L->P (in Ref. 3; BAC04004).
FT CONFLICT 259 269 AEGVLRALLAQ -> CLRTESCKAE (in Ref. 3; BAC04004).
SQ SEQUENCE 755 AA; 81488 MW; CF72D0C8EC9F69E7 CRC64;

Query Match
Best Local Similarity 27.0%; Score 107; DB 1; Length 755;
Matches 70; Conservative 23; Mismatches 106; Indels 60; Gaps 10;

QY 11 BEALETLLCYELPHHMLLELLATTT-----HCRINCGPGPAQLQALHRCPLAVCLA 65
DB 159 BEAERERQAACQCPHGEALRFLCOPCSQLCRCKLD-----PHLDHCLPLAEAVR 210
QY 66 QQLPEDEPGQSSSVFPMVKLVDSMGWELASVRR---ALCOLMD-----HEPRTG 113
DB 211 ABRRLGSG-----LLAGVDNNVVELBAKRVKELARLBAQARVGTVEEAAG 261
QY 114 VRGTGVLVFSELAFLHRSFGDLTAERKQICDPLYGRVQ-AERQALAR-----163
DB 262 VLR--ALLAQKQEVYLGQLRAHVEAAEERLAELREGVQVAAAFARVLSTGREA 319
QY 164 -----LRRTQAHSHVAFPGSCGRLTEGQDEERSTRLK---DLGRYFEESGQEPG 211
DB 320 EILSLBGLAIAQRRLQOLGCCPWAHPAPAPCLPLQLSLHPGLLDKNCHLRLSFEEOQPKDG 379
QY 212 GMEBA--QGPEPGQARLD 228
DB 380 GKXQAGTQGGESQSRRED 398

RESULT 13
Q96D36 HUMAN PRELIMINARY; PRT; 809 AA.
ID Q96D36
AC Q96D36
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DFRZP586J0619 protein (Fragment).
GN Name=DKFPZP586J0619;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP NCBIROTIDE SEQUENCE.
RC TISSUE=Lymph;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
```

```
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Huik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherich A., Schein J.E., Jones S.U.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NCBIROTIDE SEQUENCE.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013367; AAH13367.2; mRNA.
FT NON TER 1
SQ SEQUENCE 809 AA; 89682 MW; 0D557E2040458DB4 CRC64;

Query Match
Best Local Similarity 26.2%; Score 107; DB 2; Length 809;
Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

QY 29 LELLATTYHCRINCGPGPAQLQALHRCPLAVCLAQQL-----PEDPGQS--SSV 79
DB 25 LQALATV-----LSSPPGALVMSM-HRSHFLCPLRLQCOYQRCVPODTGSSFLKV 78
QY 80 EFDVVKLVDSMGWELASVRRALCOL---QMDHBRPTGVRGTGVLFVSESLAFH-----130
DB 79 LLMQLQMLDPSGVGGGLRQLMLASQASAGRLSDVRG-GILRLAELAFQDLEVS 137
QY 131 -----LRSPGDLTAERKQICDPLYGRVQAREQALRLRRTFOAFHSAV--FPSC 179
DB 138 STVRAVIATLRS-GEQGSVPEPLISKVLQGLIVRSFHELLTAPFSAADADAAPPAC 196
QY 180 GPCLEQODEERSTRKLDLGRYFEER--GQEP--GMEBA--QGPEPGQARLDQWEDQV 233
DB 197 KPVVV-----VSSL--LQEEPLAGGRGADGGLAVRRLGPSGLLVDMLEML 244
QY 234 RCDIRQPLSLRPREK-----FSSRAVARIFHGIGSPCYPAQVYQDRFMWKY-----281
DB 245 DPEV--VSSCPDLQLRLFSRR-----KKGQAQVPS-----FRPYLLTLFTHQ 286
QY 282 -----LHLSFHALVGLATER 296
DB 287 SSWPTLHQICIRVLGKSRQ 306

RESULT 14
Q6NT70 HUMAN PRELIMINARY; PRT; 900 AA.
ID Q6NT70
AC Q6NT70
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE DFRZP586J0619 protein (Fragment).
GN Name=DKFPZP586J0619;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN [1]
RP NCBIROTIDE SEQUENCE.
RC TISSUE=Brain;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
```

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Rabeck S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W., Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.E., Skalska U., Smailus D.E., Schercher A., Schin J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences.";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]

NCLEOTIDE SEQUENCE.

RP TISSUE=Brain;

RC Strausberg R.;

RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC069262; A069262.1; -; mRNA.

RT NON TER

SQ SEQUENCE 900 AA; 99543 MW; D78E27BB601BBAE2 CRC64;

Query Match 6.6%; Score 107; DB 2; Length 900;

Best Local Similarity 26.2%; Pred. No. 5.4;

Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

QY 29 LELATTYTHCRINCPGPAOLALAHRCPLAVCLAQOL-----PEDPGQGS--SSV 79

DB 116 LQALATL-----LSPHFGALVMSW-HRSHFLACPLRLQLCQYRCVPODTGFSFLFKV 169

QY 80 EEDWVKLVDSMGWELASVRLACQ--QMDHPRPGVARGGVVVESELAFH----- 130

DB 170 LLQMLQWLDSPGVREGGPIRLAQLRMLASQASARRISDVRG-GLRLAELALAFRODLEVVS 228

QY 131 -----LRSPGDLTAEEKQICDPLVGRVQAREQALRLRRTFOAFHSA--FPSC 179

DB 229 STVRAVAVITLRS-GEQCSVEPDLISKVLQGLIENRSPHLEILLTAFFSATDAASPFPAC 287

QY 180 GPCLEQODEERSTRKDLIGRYFEEER--GOEP--GEMEDAQEPGQARLDQWEDQV 233

DB 288 KPVVV-----VSSIL--LQEEEPPLAGKPGADGSLDAVRLGPSSGLLVWLEML 335

QY 234 RCDIRQFSLRPEEK-----FSSRAVARIFHGIGSCPYAQVYGGQRRFWRKY----- 281

DB 336 DPEV--VSSCPDLQRLILFSRR-----KKGQAQVPS-----FRPYLLTLFTHQ 377

QY 282 -----LHLSFHALVGLATEE 296

DB 378 SSMPTLHQCIRVILGKSRQ 397

RESULT 15

OGUX74 HUMAN PRELIMINARY; PRT; 1029 AA.

ID OGUX74_

AC OGUX74_

DT 05-JUL-2004 (TREMREL. 27, Created)

DT 05-JUL-2004 (TREMREL. 27, Last sequence update)

DT 05-JUL-2004 (TREMREL. 27, Last annotation update)

DE HL1V1821.

GN ORFNames=UNQ1821;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.

OC NCB1_TaxID=9606;

OX NCB1_TaxID=9606;

RP NCLEOTIDE SEQUENCE.

RX MEDLINE=22887296; PubMed=1295309; DOI=10.1101/gr.1293003; Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J., Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,

Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J., Seethagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A., Vanden R.L., Watanabe C., Wland D., Woods K., Xie M.-H., Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D., Wood W.T., Godowski P.J., Gray A.M.; "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";

RT Genome Res. 13:2265-2270(2003).

DR EMBL: AY358482; AA088486.1; -; mRNA.

SQ SEQUENCE 1029 AA; 114212 MW; D28FD029DD0A797 CRC64;

Query Match 6.6%; Score 107; DB 2; Length 1029;

Best Local Similarity 26.2%; Pred. No. 6.3;

Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

QY 29 LELATTYTHCRINCPGPAOLALAHRCPLAVCLAQOL-----PEDPGQGS--SSV 79

DB 245 LQALATL-----LSPHFGALVMSW-HRSHFLACPLRLQLCQYRCVPODTGFSFLFKV 298

QY 80 EEDWVKLVDSMGWELASVRLACQ--QMDHPRPGVARGGVVVESELAFH----- 130

DB 299 LLQMLQWLDSPGVREGGPIRLAQLRMLASQASARRISDVRG-GLRLAELALAFRODLEVVS 357

QY 131 -----LRSPGDLTAEEKQICDPLVGRVQAREQALRLRRTFOAFHSA--FPSC 179

DB 358 STVRAVAVITLRS-GEQCSVEPDLISKVLQGLIENRSPHLEILLTAFFSATDAASPFPAC 416

QY 180 GPCLEQODEERSTRKDLIGRYFEEER--GOEP--GEMEDAQEPGQARLDQWEDQV 233

DB 417 KPVVV-----VSSIL--LQEEEPPLAGKPGADGSLDAVRLGPSSGLLVWLEML 464

QY 234 RCDIRQFSLRPEEK-----FSSRAVARIFHGIGSCPYAQVYGGQRRFWRKY----- 281

DB 465 DPEV--VSSCPDLQRLILFSRR-----KKGQAQVPS-----FRPYLLTLFTHQ 506

QY 282 -----LHLSFHALVGLATEE 296

DB 507 SSMPTLHQCIRVILGKSRQ 526

Search completed: December 27, 2005, 21:53:12

Job time : 61.8 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:43:16 ; Search time 13 Seconds
(without alignments)
1920.620 Million cell updates/sec

Title: US-09-889-325-4_COPY_907_1208

Perfect score: 1 QLTVALDMPREAIETLLCY.....HLSPHALVGLATELLQVAR 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 segs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents AA:*
2: /cgn2_6/prodata/1/1aa/5_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/H_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/RE_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 1617 | 100.0 | 1208 | 2 | US-09-463-702A-2 |
| 2 | 1617 | 100.0 | 1208 | 2 | US-09-699-135-2 |
| 3 | 108.5 | 6.7 | 762 | 2 | US-09-252-991A-29423 |
| 4 | 107 | 6.6 | 1029 | 2 | US-09-829-769-7 |
| 5 | 107 | 6.6 | 1029 | 2 | US-10-033-301-22 |
| 6 | 102.5 | 6.3 | 279 | 2 | US-09-252-991A-18933 |
| 7 | 102 | 6.3 | 817 | 2 | US-09-252-991A-25887 |
| 8 | 99.5 | 6.2 | 512 | 2 | US-09-252-991A-22339 |
| 9 | 94.5 | 5.8 | 288 | 2 | US-09-127-928-2 |
| 10 | 94.5 | 5.8 | 1055 | 2 | US-09-502-540-15453 |
| 11 | 92.5 | 5.7 | 515 | 2 | US-09-252-991A-19259 |
| 12 | 92 | 5.7 | 416 | 2 | US-09-808-387-38 |
| 13 | 91.5 | 5.7 | 2548 | 2 | US-09-172-422-1 |
| 14 | 90 | 5.6 | 490 | 2 | US-09-252-991A-18962 |
| 15 | 89.5 | 5.5 | 565 | 2 | US-09-008-481A-6 |
| 16 | 89.5 | 5.5 | 565 | 2 | US-09-195-666A-5 |
| 17 | 89.5 | 5.5 | 565 | 2 | US-09-195-666A-49 |
| 18 | 89.5 | 5.5 | 565 | 2 | US-09-309-592-6 |
| 19 | 89.5 | 5.5 | 565 | 2 | US-09-635-705-5 |
| 20 | 89.5 | 5.5 | 565 | 2 | US-09-635-705-49 |
| 21 | 89.5 | 5.5 | 565 | 2 | US-09-634-858A-5 |
| 22 | 89.5 | 5.5 | 565 | 2 | US-09-634-858A-49 |
| 23 | 89.5 | 5.5 | 565 | 2 | US-08-869-927C-5 |
| 24 | 89.5 | 5.5 | 565 | 2 | US-08-869-927C-49 |
| 25 | 89 | 5.5 | 1012 | 2 | US-08-811-481-16 |
| 26 | 89 | 5.5 | 1012 | 2 | US-09-876-527-16 |
| 27 | 88.5 | 5.5 | 600 | 2 | US-09-252-991A-16741 |

| | | | | | | |
|----|------|-----|------|---|----------------------|-------------------|
| 28 | 88 | 5.4 | 274 | 2 | US-09-252-991A-27466 | Sequence 27466, A |
| 29 | 88 | 5.4 | 3816 | 2 | US-09-428-517-3 | Sequence 3, App11 |
| 30 | 87.5 | 5.4 | 518 | 2 | US-09-252-991A-27664 | Sequence 27664, A |
| 31 | 86.5 | 5.3 | 290 | 5 | US-09-252-991A-27484 | Sequence 27484, A |
| 32 | 86.5 | 5.3 | 915 | 2 | US-09-252-991A-23637 | Sequence 23637, A |
| 33 | 86 | 5.3 | 998 | 2 | US-09-252-991A-24402 | Sequence 24402, A |
| 34 | 86 | 5.3 | 1735 | 2 | US-09-902-540-14547 | Sequence 14547, A |
| 35 | 86 | 5.3 | 1747 | 2 | US-09-902-540-14765 | Sequence 14765, A |
| 36 | 85 | 5.3 | 989 | 2 | US-09-199-637A-273 | Sequence 273, App |
| 37 | 85 | 5.3 | 1190 | 2 | US-09-902-540-12293 | Sequence 12293, A |
| 38 | 85 | 5.3 | 1673 | 2 | US-09-418-710-70 | Sequence 70, App1 |
| 39 | 85 | 5.3 | 1673 | 2 | US-09-839-479-69 | Sequence 69, App1 |
| 40 | 85 | 5.3 | 1674 | 2 | US-09-418-710-1 | Sequence 1, App11 |
| 41 | 85 | 5.3 | 1674 | 2 | US-09-839-479-1 | Sequence 20, App1 |
| 42 | 85 | 5.3 | 1681 | 2 | US-10-037-417-20 | Sequence 68, App1 |
| 43 | 85 | 5.3 | 1697 | 2 | US-10-037-417-68 | Sequence 28824, A |
| 44 | 84.5 | 5.2 | 457 | 2 | US-09-252-991A-28824 | Sequence 21993, A |
| 45 | 84.5 | 5.2 | 709 | 2 | US-09-252-991A-21993 | Sequence 21993, A |

ALIGNMENTS

```
RESULT 1
US-09-463-702A-2
; Sequence 2, Application US/09463702A
; Patent No. 6335435
; GENERAL INFORMATION:
; APPLICANT: AGENE Research Institute, Co., Ltd.
; APPLICANT: HIRAKI AND ASSOCIATES
; APPLICANT: SHIMAMOTO, AKIRO
; APPLICANT: KITAO, SAORI
; APPLICANT: FURUICHI, YASUHIRO
; TITLE OF INVENTION: HUMAN GENE RECO4 ENCODING HELICASE
; FILE REFERENCE: HIRALISO
; CURRENT APPLICATION NUMBER: US/09/463,702A
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/JP98/03114
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: JAPAN 9/200387
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-463-702A-2

Query Match      100.0%; Score 1617; DB 2; Length 1208;
Best Local Similarity 100.0%; Pred. No. 2.4e-178;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QLTVALDMPREAIETLLCYELAPHHMELATTYTHCRNLCPGGPAQLALAHRCPEL 60
Db 907 QLTVALDMPREAIETLLCYELAPHHMELATTYTHCRNLCPGGPAQLALAHRCPEL 966

61 AYLCAQLPDPDQGGSSVEFPMVKLVDSMGWELASVPRALCOLQWDHPRFGVGRGTGV 120
QY 907 QLTVALDMPREAIETLLCYELAPHHMELATTYTHCRNLCPGGPAQLALAHRCPEL 966

967 AYLCAQLPDPDQGGSSVEFPMVKLVDSMGWELASVPRALCOLQWDHPRFGVGRGTGV 1026
Db 967 AYLCAQLPDPDQGGSSVEFPMVKLVDSMGWELASVPRALCOLQWDHPRFGVGRGTGV 1026

121 LVEFSLAFHLSPGDLTAEBKQICDFLYGVRQERQALRLRTFOAFHSVAFPSGC 180
QY 121 LVEFSLAFHLSPGDLTAEBKQICDFLYGVRQERQALRLRTFOAFHSVAFPSGC 180

1027 LVEFSLAFHLSPGDLTAEBKQICDFLYGVRQERQALRLRTFOAFHSVAFPSGC 1086
Db 1027 LVEFSLAFHLSPGDLTAEBKQICDFLYGVRQERQALRLRTFOAFHSVAFPSGC 1086

181 PCLBQDEERSRLKDLGRYFEEBEGQBPGEAQAQBPQARIQDMEQVRCDIQF 240
QY 181 PCLBQDEERSRLKDLGRYFEEBEGQBPGEAQAQBPQARIQDMEQVRCDIQF 240

1087 PCLBQDEERSRLKDLGRYFEEBEGQBPGEAQAQBPQARIQDMEQVRCDIQF 1146
Db 1087 PCLBQDEERSRLKDLGRYFEEBEGQBPGEAQAQBPQARIQDMEQVRCDIQF 1146

241 LSLRPEKSSRAVARIIFGIGSPCYPAOVYQODRRFMKXYLHLSFHALVGLATELLQV 300
QY 241 LSLRPEKSSRAVARIIFGIGSPCYPAOVYQODRRFMKXYLHLSFHALVGLATELLQV 300

1147 LSLRPEKSSRAVARIIFGIGSPCYPAOVYQODRRFMKXYLHLSFHALVGLATELLQV 1206
Db 1147 LSLRPEKSSRAVARIIFGIGSPCYPAOVYQODRRFMKXYLHLSFHALVGLATELLQV 1206
```

QY 301 AR 302
Db 1207 AR 1208

RESULT 2

US-09-699-135-2
; Sequence 2, Application US/09699135
; Patent No. 6472513
; GENERAL INFORMATION:
; APPLICANT: AGENE Research Institute, Co., Ltd.
; APPLICANT: HIRAKI AND ASSOCIATES
; APPLICANT: SHIMAMOTO, AKIRO
; APPLICANT: KITAO, SAORI
; APPLICANT: FURUICHI, YASUHIRO
; TITLE OF INVENTION: HUMAN GENE RECO4 ENCODING HELICASE
; FILE REFERENCE: HIRAI150
; CURRENT APPLICATION NUMBER: US/09/699,135
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US/09/463,702A
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/JP98/03114
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: JAPAN 9/200387
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-699-135-2

Query Match 100.0%; Score 1617; DB 2; Length 1208;
Best Local Similarity 100.0%; Pred. No. 2,4e-178;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLTVALDMPREAIETLLCYELHPHMLLATTYTCRANCRCGPAQLALHRCPL 60
Db 907 QLTVALDMPREAIETLLCYELHPHMLLATTYTCRANCRCGPAQLALHRCPL 966
QY 61 AVCLAQQLPEDPGCGSSSVFEDMYKLVDMSGMEIASVRLCOLQMDHEPTGVRRGTGV 120
Db 967 AVCLAQQLPEDPGCGSSSVFEDMYKLVDMSGMEIASVRLCOLQMDHEPTGVRRGTGV 1026
QY 121 LVFSESLAFHRSFGDLTAEEKDQICDPLVGRVQAREQALARLRTFQAFHSVAFPSCG 180
Db 1027 LVFSESLAFHRSFGDLTAEEKDQICDPLVGRVQAREQALARLRTFQAFHSVAFPSCG 1086
QY 181 PCLEQODEERSTRKDLGRYFEEEGOEPCGMDACQPEFGARLQDMEQVRCDIROF 240
Db 1087 PCLEQODEERSTRKDLGRYFEEEGOEPCGMDACQPEFGARLQDMEQVRCDIROF 1146
QY 241 LSLRPEKFSRAVARIFHGIGSPCPYAOVYGODRRFMRKYLHSPALVGLATEBLLQV 300
Db 1147 LSLRPEKFSRAVARIFHGIGSPCPYAOVYGODRRFMRKYLHSPALVGLATEBLLQV 2006
QY 301 AR 302
Db 1207 AR 1208

RESULT 3
US-09-252-991A-29423
; Sequence 29423, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29423
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29423

Query Match 6.7%; Score 108.5; DB 2; Length 762;
Best Local Similarity 23.9%; Pred. No. 0.00311; Indels 129; Gaps 14;
Matches 72; Conservative 21; Mismatches 79;

QY 46 GPAQLALAH---RCPLAVCLAQQLPEDPGCGSSSVFEDMYKLVDMSGMEIASVRLALC 102
Db 470 GPSLQGPRRHPRRRRPLA-----ETPGQRRSAEH-----APGTGSD----- 507
QY 103 QLOMDHEPTGVRRGTGVLVFSSELAFLHRSFGDLTAEEKDQICDPLVGRVQAREQA-- 160
Db 508 ---DHGPRDPGR-----DRPGELVQSRPRRSCAAGQGVHRRRLQADG 547
QY 161 -----LARIPTFOAFHSVAFPSCGRCLEQODEERSTRKDLGRYFEEEGOE 210
Db 548 GRPATAGRLPLRLK-----RPGPAHLPSFTER--RLAQVLGMRWRDEPRRAP 594
QY 211 GGMEDA-----QGP-----EPGOARLQDMEQV 233
Db 595 GGRHARBDGHPRRQAAGMGLRCFVPHQGGGQDARGOALQLRPGRLRLDGRV 654
QY 234 RCDIRQLSLRPEKFSRAVARIFHG-----IGSPCY-----PAQVQODR 275
Db 655 QDQLPRRIP-RAURRRGRRAAPASAGEPRPATDQPHRHHRVGDPHHFGGRPA---GQPR 710
QY 276 R 276
Db 711 R 711

RESULT 4

US-09-929-769-7
; Sequence 7, Application US/09929769
; Patent No. 6914130
; GENERAL INFORMATION:
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Polakis, Paul
; APPLICANT: Shou, Jianyong
; APPLICANT: Smith, Victoria
; APPLICANT: Soriano, Robert
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P5007R1-US
; CURRENT APPLICATION NUMBER: US/09/929,769
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/089,653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/090,355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/104,257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/141,037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: PCT/US99/12252

```

; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/20111
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: PCT/US00/00376
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PCT/US00/04342
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/13705
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/20118
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/888,257
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-929-769-7

```

```

Query Match      6.6%; Score 107; DB 2; Length 1029;
Best Local Similarity 26.2%; Pred. No. 0.0072;
Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

```

```

QY 29 LELATTYTHCHLNCGPGPAOLALAHRCPLAVCIAOUL-----PEDPGQS--SSV 79
DB 245 LQALMTL-----LSSPHGALVMSH-HRSHPLACPLRLQCYQKCVPDGTGFSLSFLKV 298
QY 80 BPDWVKLVDSMGWELASVRRALCOL--QMDHPRRTGVRGTVLVEFSELAFH----- 130
DB 239 ILQMTLQWLDSPVEGPGPLRAQLRMLASQASAGRRLSDVWG--GLRLAELALAFODLEVVS 357
QY 131 -----LRSRGLTAEKXQICDPLXGRVQARERQALRLRTFQAFHVA--FPSC 179
DB 358 STVRVAVIATLRS-GEQCSVEPDLISKVLQGLIEVNSPHLEILLTAFSATPADASPPFAC 416
QY 180 GPCLSEQDEBERSTRLLKDLIGRYFEERE--GOBP--GMEDAQGPEPQARLQDMEDQV 233
DB 417 KVVVV-----VSSLL--LQEBEPLAGKPGADGSLAVRLGSPSSGLVDMLEML 464
QY 234 RCDIROFTSLRBEK-----FSSRAVARIFHGIGSPCYPAQVYGQDERRFRKTY----- 281
DB 465 DEEV--VSSCPDLQRLILFSRR-----KKGQAQVPS-----FRPYLLTLFTHQ 506
QY 282 -----LHLSFHLVGLATEE 296
DB 507 SSMPTLHOCIRVLLGKSRQ 526

```

```

RESULT 5
US-10-033-301-22
; Sequence 22, Application US/10033301
; Patent No. 6930172
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Guirey, Austin L.

```

```

; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2930R1C6
; CURRENT APPLICATION NUMBER: US/10/033,301
; CURRENT FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 22
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-033-301-22

```

```

Query Match      6.6%; Score 107; DB 2; Length 1029;
Best Local Similarity 26.2%; Pred. No. 0.0072;
Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

```

```

QY 29 LELATTYTHCHLNCGPGPAOLALAHRCPLAVCIAOUL-----PEDPGQS--SSV 79
DB 245 LQALMTL-----LSSPHGALVMSH-HRSHPLACPLRLQCYQKCVPDGTGFSLSFLKV 298
QY 80 BPDWVKLVDSMGWELASVRRALCOL--QMDHPRRTGVRGTVLVEFSELAFH----- 130

```

```
Db 299 LKQLMLWDSGCVGGLPRAQLMFLASQASGRSLSDVRG--GLRLMALFRDDELYVS 357
QY 131 -----LRSPDDLTAEEKDQICDPLYGVARERQALRLRTTFOAFHSVA--FPSC 179
Db 358 STVAVATATRS--GEOSVPEPDLISKVLQGLIEVRSPLIEBLITAFSATDAASPFPAC 416
QY 180 GPCLEQODEERSTRLKDLGRYFEEB---GQEP---GGMEDAGPEPGQARLQDMEQV 233
Db 417 KPVVV-----VSSL--LQEEPLAGKFGADGSGLEAVRLGSPGLVDMLEML 464
QY 234 RCDIRQFLSLRPEEK---FSSRAVARLFGIGSPCYPAQVYGQDRFRFRKY----- 281
Db 465 DPEV---VSSCPDLQLRLPFSR-----KKGQAQVPS-----FRPYLLTLPTHQ 506
QY 282 -----LHLSFHALVGLATEE 296
Db 507 SSWPTLHQICIRVLIGKSRQ 526
```

RESULT 6

```
US-09-252-991A-18933
; Sequence 18933, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18933
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18933
```

```
Query Match 6.3%; Score 102.5; DB 2; Length 279;
Best Local Similarity 25.8%; Pred. No. 0.0034;
Matches 64; Conservative 35; Mismatches 106; Indels 43; Gaps 11;
```

```
QY 59 PLAYCLAQQLPEBDGSSSVFPMVTLVDSGMELASV---RRALCOLQMDHEPRTGVR 115
Db 5 PLPV--EQLLVAAHQGHRIEFD--QRADTLGGDDADRIDDRQPEPELQADAQ----- 55
QY 116 RGTGVVVEPSLAFHLSFGDLTAEEKDQICD--FLYGRVQAREQALRLRTTFOAFHS 173
Db 56 ----LADVAF--BHVHPEBDHPQSRREQHLDROQRHRECCRAKQAPABEKKAEQGED 108
QY 174 VAFPSGCPGLEQODEERS--TRLKDL-----GRYEBEESGPGGMBDAG 218
Db 109 DG--EIDPGEYQYDRQADAREADLLQOVGLDEHALAARDPFEQAPGQAGAEITAVG 166
QY 219 P---EPQARLQDM--EDQVRCDIRQFLSLRPEEKFSRAVARLFGIGSPCYPAQVYGQ 273
Db 167 QAVVDPRQARAHNLGEBHRVDDHRRQRIENRQFAEGQPAVARLEMLDAENENTTAPQ 226
QY 274 DRFRFRKY 281
Db 227 GTRQAKKH 234
```

RESULT 7

```
US-09-252-991A-25887
; Sequence 25887, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25887
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25887
```

```
Query Match 6.3%; Score 102; DB 2; Length 817;
Best Local Similarity 28.6%; Pred. No. 0.019;
Matches 67; Conservative 23; Mismatches 62; Indels 82; Gaps 16;
```

```
QY 49 QLOALHRCPLAYCLAQQLPEBDGSSSVFDP-----MVKLVDSGMELASV----- 97
Db 299 EAQRRIHR--PLAVHIL--VPVDFGQAGFPRAESDRQFLAARLIDRM--QDAQVFDDRLA 352
QY 98 -----RRALCOLQMDHEPRTGVR-----GTGVL-----VESSELAFHLS- 133
Db 353 FRGVGQRAQRLAGQFLAAH--PRRGVQRAATVAEDGAGLVRRQHVGEPRPAFHRBDP 411
QY 134 -----PGDLTAEEKDQICDPLYGVARERQALRLRTTFOAFHSVAFPSG 180
Db 412 ELATFRTRRSMGCD--AAREQPAQ--GGRDQRHQQRHQRHQRQ-----VAAGEQG 460
QY 181 PCLE---QODEERSTRLKDLGRYFEEBGEQEPGKEDA---GQEPQARL 226
Db 461 ERLEGHNYQODEQDQADQOD--VQRHF-----VGGILPFGAFHQGDHPVQGRIL 506
```

RESULT 8

```
US-09-252-991A-22339
; Sequence 22339, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22339
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22339
```

```
Query Match 6.2%; Score 99.5; DB 2; Length 512;
Best Local Similarity 24.1%; Pred. No. 0.019;
Matches 63; Conservative 23; Mismatches 80; Indels 89; Gaps 12;
```

```
QY 108 HEPRTGVRGTVLVEPSLAFHLSFGDLTAEEKDQICDPLYGVARERQALRLRTT 167
Db 99 HEVLAVGYQR--LAVE-----HADAPVVLGVE-----LIGQQQDRALQGLARRRQ 142
QY 168 FOAF-----HSAVPSGCPGLEQODE-----RSTRKDLGRYFE-----EESQ-- 208
Db 143 FAVVVDLHATGEMALGNLQHOREAQAPGHPRLPYLRAFLVEBDIGERRHQLPATYQVQVD 202
QY 209 -----EPGMDAGPEPGQARLQDMEQVQCD- 236
```

Db 203 LVGAODRCRVVHHHQAALGLSALIGMVDPGGLADQCGVEIGQARVVLASDQFDVDA 262
Qy 237 -----IRQFLSLRPEEK-----SSRAVA-----RIFHIGISPCYPAQVYGODRRF-- 277
Db 263 EAGADPYEVGQLRKRGRFRVGVIEDRHVVAChPRTTRHAPGAETVEEDILLVA 322
Qy 278 WRKYLHLSFHAL---VGLATE 295
Db 323 AQAVQHARAHLMDPALLATE 343

RESULT 9
US-09-127-928-2
Sequence 2, Application US/09127928
Patent No. 6472172
GENERAL INFORMATION:
APPLICANT: Deng, Gang
APPLICANT: Lin, Jiling-Huey
APPLICANT: Moreser, Michael J
TITLE OF INVENTION: DNA Encoding a No. 6472172el Human Inhibitor-of-Apoptosis
TITLE OF INVENTION: Protein
FILE REFERENCE: DNA Encoding H1AP3
CURRENT APPLICATION NUMBER: US/09/127,928
CURRENT FILING DATE: 1998-07-31
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 298
TYPE: PRT
ORGANISM: Homo sapiens
US-09-127-928-2

Query Match 5.8%; Score 94.5; DB 2; Length 298;
Best Local Similarity 20.8%; Pred. No. 0.032;
Matches 62; Conservative 32; Mismatches 113; Indels 91; Gaps 12;

Qy 25 PHHWLELATYTHRLNCPGPAQLQALHRCPLAVCLAQQLPEDPGQSSSVFPMV 84
Db 16 PSHMA-----AGDPT-----EKGRSL-----GSPVGLDTC 45
Qy 85 KLVDMSGWEIASVRALCOLQMDHBRPTGVRGTG-----LVESSELAFLRSPG 135
Db 46 RAMDVDDQILQLPLRTESEEBGAGATLSRPAFPGSGSELRLASFYDPLTAEVPP 105
Qy 136 DITAE-----EKQOI-CDFLYGRVQARER-----QALARK-----RT 167
Db 106 BILAAAGFHTHGODKVRFCFCYGGIQSWKRGDDPWTEHAKMPPSCQFLNRSGKRDVHS 165
Qy 168 FGAFTSVAFPSGQCLQODERSTRLKDLG-----RYFEEBEGEPGMDAQP 219
Db 166 VERTHSQILGSDWPMEEPDAAPVADSVASGYPELPTRRREVOSASQEPGVSPEAQ 225
Qy 220 -----EPQARLQDMDQVRCDIROFLSLRPEEKSSRAVAIF-----HGISPCYP 267
Db 226 RAMWVLEPFGAR-----DVEAQRLRLOBERCKVCLDAVSLVFPCGHVLCABCAP 277

RESULT 10
US-09-902-540-15453
Sequence 15453, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 15453
LENGTH: 1055
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-15453

Query Match 5.8%; Score 94.5; DB 2; Length 1055;
Best Local Similarity 26.3%; Pred. No. 0.21;
Matches 82; Conservative 40; Mismatches 103; Indels 87; Gaps 20;

Qy 7 LDMPEBAIR-----TLTCYLEHHPHMLLATYTHRLNCPG-----GPA 48
Db 344 LDNLEQAVOHLPAVTLGWLQALPR--ARFLATSR--ALGLPGRLLDLAPLSVEEES 399
Qy 49 QLOALARCPLAVCL-AQQLPEDPG-----QSSSVFPMVKLV--SMGWEIASVRA 100
Db 400 RLEELVRSDD--AVRLAQRAREARGIFELTGAAPVADIVRODDGIALAIELAAAMA 456
Qy 101 LCOLQMDHBRPTGVRGTGVLVEFSELAFLRSPGDLTAEEKQICDFLYGRVQARERQA 160
Db 457 LLSVQLRERLT--R-----FELLRKGRPDGRARQATLRGALDMSWMLRPEERT 506
Qy 161 LARLR-----RTFOAFHSV-AFPGCGPCLQODERSTRLKDLGRYFEEBEGEPGME 214
Db 507 LARCSVFEGFTLRAAEAVILGLPDPGPAV--LDVLQSLRVSIL-RVLEAE--LPG-- 557
Qy 215 DAGPBPQCARLQDMDQVRCDIROFLSLRPEEKSSRAVAIFHIGISPCYPAQVYGOD 274
Db 558 -----GDSRLQGYE-----SIROYAARLAEB-----GVGNAAALARRHAD- 593
Qy 275 RRFWRKYLHLSF 286
Db 594 ---W--YLSLAY 600

RESULT 11
US-09-252-991A-19259
Sequence 19259, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19259
LENGTH: 515
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19259

Query Match 5.7%; Score 92.5; DB 2; Length 515;
Best Local Similarity 23.2%; Pred. No. 0.12;
Matches 69; Conservative 40; Mismatches 105; Indels 83; Gaps 15;

Qy 48 AQOLALA--HRCPLAVCLAQQLPEDPGQSSSVFPMVKLVDSMGWEIASVRR----- 99
Db 96 AQOLALVGGRRRALQAQVGHVLD-----QLVEVLDIVYHALQVLAQAVGIG 144
Qy 100 ALCOLQMDHBRPTGVRGTGVLVEFSELAFLRSPGDLTAEEKQICDFLYGRVQARERQ 159
Db 145 GLDHLQAEAGPRH--RRALQVGHANHLALHRQGVVDVIGH-----MVEBGAQASHRV 195
Qy 160 ALARLRRTFOAFHSVAFPSGCL-----EQDERSTRLKD-----LGR 200
Db 196 ATAHLDASTIQ--AAGDPGRGLQAALLOTTHQVDDQADEAQARAEQOQQLRRIGV 253

Query 201 YFEEB-EGOEFGMEDAOGPEG-----QARLODMEDOVRC 236
Db 254 HLEGRAGLOHFGREDA-GEDADRVAVLAQRHNRVALHHSBALVVEVQLDLE-QVHRE 311
Query 237 IROFLSRPEEKFSRAVARIFHG-IGSPCYAQVYGODRRFWRKYLHSF-HALVG 291
Db 312 AEAIVADDP-----RKVVRFLFAGRVAHQFVHQVHGRFQLDLHLHAEHQVLG 362

RESULT 12
US-09-808-387-38
; Sequence 38, Application US/09808387
; Patent No. 6962985
; GENERAL INFORMATION:
; APPLICANT: Kaia Palm
; APPLICANT: Toni's Timmusk
; APPLICANT: Cemines Research
; TITLE OF INVENTION: MAMMALIAN NEURALIZED FAMILY OF
; TITLE OF INVENTION: TRANSCRIPTION REGULATORS AND USES THEREFOR
; FILE REFERENCE: CEMRES.001A
; CURRENT APPLICATION NUMBER: US/09/808,387
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 416
; TYPE: PR
; ORGANISM: Homo sapien
US-09-808-387-38

Query Match 5.7%; Score 92; DB 2; Length 416;
Best Local Similarity 18.6%; Pred. No. 0.1;
Matches 54; Conservative 41; Mismatches 90; Indels 106; Gaps 12;

Query 22 ELHPHMLLEL-----ATTYHCLNCPGSPAOLOALAHRCPP-----AVCLAQ 67
Db 141 EMKPHSYLDIIRSGLDVVEASSYSNEQOLCPYAAAGCRGDAQVYLHGVCEICHLQV 200
Query 68 L-PEDPGOGSSSVFDMVKLVDSMGELASVRLCOLQMDHERPTGVRGTGVLEFSE 126
Db 201 LHPPDP-----BQRKAHERICMLTFEHE-----ME 225

Query 127 LAFLRLPGDLTAEKDOIC-----DPLYGRVQAREO-----ALARLRPTQ 169
Db 226 KAPFQA-----SQDKVCSICMEVILEKASASERRGILSNCHTYCISCTIQMKAE 278

Query 170 AFHSVAPSCGPC-----LEQDE--ERSTRLDLIG-----RYFEEBEGQEP 210
Db 279 QFENPIIKSCPECKVISEFVIPSYYWVEDQKKNELIAPFQMGKKACKYFEGQKGTCP 338

Query 211 GGMEDAGPEPGCARLODMEDOVRCDIRQFLSLRPEEKFSRAVARIFHGT 261
Db 339 FGSKCLYRHAYPPDRLAEP-----KPRKLSQGTVRFPNSV 376

RESULT 13
US-09-172-422-1
; Sequence 1, Application US/09172422A
; Patent No. 6300485
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen B.
; APPLICANT: Chiu, Choi Yung
; APPLICANT: Duhl, David
; APPLICANT: Gortman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNCG-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; FILE REFERENCE: 200130.442
; CURRENT APPLICATION NUMBER: US/09/172.422A
; CURRENT FILING DATE: 1998-10-14

Query 47 PAQOLAHRCPPAVCLAQQLPEDRG-----QSSSVFDMVKLVDSM 90
Db 1798 PDELAAY-HTPPLIS-----PELPSCKEKFEKNEKPSPKARKKSVKISVAL-DSM 1848

Query 91 GWEIASVRLCOLQMDHERPTGVRGTGVLEFSEL-----AFHLRPGDLTAE--KDQ 144
Db 1849 HMQNDSTVQ-----IIASVSDKSMDFELKRYNDLNDSDSKDT 1887

Query 145 ICDFLYGRVQAREKQALARTFQAHSVAPSCGPLEQODEERSTRLDLGRYFEE 204
Db 1888 LVDPVF-----KKALKERQINFSEYSSAL-----AMDGKISRYKDLVALFEQI 1932

Query 205 BEGGEFGMEDAOGPEPGCARLODMEDOVRCDIRQFLSLRPEEKFSRAVARIFHGTGSP 264
Db 1933 LEKTRLEQSDSGESP-----VRVWNTFKVFLDEYWN-----EFKT-----SD 1972

Query 265 CYPAQVYGODRRFWRK 280
Db 1973 CTARKVPTKTKRK 1988

RESULT 14
US-09-252-991A-18962
; Sequence 18962, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18962
; LENGTH: 490
; TYPE: PR
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18962

Query Match 5.6%; Score 90; DB 2; Length 490;
Best Local Similarity 25.1%; Pred. No. 0.2;
Matches 76; Conservative 35; Mismatches 108; Indels 84; Gaps 18;

Query 38 HCRNLCPGSPAOLOALAHRC-----PDLAVCLAQQL-----PEDPGOGSSSVFDMVKL 86
Db 24 HLFVECPA--RQLOQLHHRIDIPALAEKRLAQLRLEGFLAFROPGEAAAFGVPTLQA 81

Query 87 VD-SMG--WEIASVRLCOLQMDHERPTGVRGTGVLEFSELAFLRLRPGDLTAEKD 143
Db 82 EDWTLGGVGLATYARAV-----GHOPRPRRGG-GIRLAABLA-----GNAGMFEERD 133

Query 144 QICDFLYGRVQAREKQALARTFQAHSVAPSCGPLEQO-----DEERST 192
Db 134 VLATL-----ADRRQQL-----GDVQAHQVLEASGTCFGQVGLGRGNHPQVDVPTLV 184

Query 193 RLK-----DLGRY-----FEEBEGFGME--DQGPFGCARLODMEDOV 233

Db 185 RAEPQLLLDQAQQLDLRGRHAFDLVEBQGAAMGLELADALRAGE-RTALMTEQL 243
QY 224 RCDIRPFLSLREEKFSRAVARIRFHIGSPCYPAQVYGQDRFRKTLHLSPHALVGLA 293
Db 244 A-LEQLLGNR-----RAVGRHRLVGAARPEVQAAGN-----QFLAAGLA 283
QY 294 TEE 296
|:|
Db 284 TDQ 286

RESULT 15

US-09-008-481A-6
; Sequence 6, Application US/09008481A
; Patent No. 6087151
; GENERAL INFORMATION:
; APPLICANT: ARIO, Takeshi
; APPLICANT: TANAI, Madoka
; APPLICANT: TORIGOE, Kakuji
; APPLICANT: KIRIMOTO, Masashi
; TITLE OF INVENTION: DNA CODING FOR MAMMALIAN L-
; TITLE OF INVENTION: ASPRAGINASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/008,481A
; FILING DATE: 16-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/598,369
; FILING DATE: 08-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 95-42564
; FILING DATE: 08-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,981
; REFERENCE/DOCKET NUMBER: ARIO=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 565 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-008-481A-6

Query Match 5.5%; Score 89.5; DB 2; Length 565;

Best local Similarity 26.0%; Pred. No.0.32; Mismatches 93; Indels 69; Gaps 13;

Matches 65; Conservative 23; Mismatches 93; Indels 69; Gaps 13;
QY 4 VQALDMPERAIETLLCYL-----ELHPHMLELA-----TYYHGRINCPG---- 45
Db 318 VSGIDMTSEALAKLSTVGLPELISLRQELAKDLRGEMTLPTADLH-QSSPPESTLG 376
QY 46 -GPAQLOALAHRC-----PPLAVCLAQQLPEDPGQSSSVFEDMVNKLVDMSGWE 93
Db 377 QGVARLFSL-FGQOEEDSVODAVMPSIALALAHAGLEALQALMEIGSDI-RLKDSNGQT 434
QY 94 IASVRRALCOLQMDHEPRGTGVRGTGVLVFSSEIAFHLRSPGDLTAEEKDQICDFLYGRV 153

Db 435 LIHYAA-----RNGRDGV-----VTMLIHRGMVDVNAARDGLSPLLLA-V 473
QY 154 QAREKQALARLRRTTQAFHVSVAFPSCGPELEQOD-BERSTRKLKDLGRYFEEEGQOEPBG 212
Db 474 QGRHRECIIRLIRK-----AGACLSPOLKDKAGTELICLASH--ADMEGLOQMWG 519
QY 213 MEDAQPEPG 222
|:|
Db 520 QAGADLQQPG 529

Search completed: December 27, 2005, 21:55:17
Job time : 14 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2005, 21:53:22 ; Search time 44.6 Seconds
(without alignments)
2829.248 Million cell updates/sec

Title: US-09-889-325-4_COPY_907_1208
Perfect score: 1617
Sequence: 1 QLTVALDMPERAEITLLCY.....HLSPHALVGLATELLQVAR 302

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|----------------------|
| 1 | 1617 | 100.0 | 1208 | 3 | US-09-889-325-4 |
| 2 | 1617 | 100.0 | 1208 | 5 | US-10-723-860-4382 |
| 3 | 1617 | 100.0 | 1208 | 5 | US-10-882-405-4 |
| 4 | 1617 | 100.0 | 1208 | 5 | US-10-756-149-5768 |
| 5 | 1144.5 | 70.8 | 1216 | 6 | US-11-084-955-2 |
| 6 | 427 | 26.4 | 1579 | 6 | US-11-097-143-16743 |
| 7 | 169 | 10.5 | 874 | 4 | US-10-437-963-121617 |
| 8 | 145.5 | 9.0 | 206 | 4 | US-10-425-114-68934 |
| 9 | 145.5 | 9.0 | 215 | 4 | US-10-425-115-338629 |
| 10 | 107 | 6.6 | 1029 | 4 | US-10-476-924-22 |
| 11 | 107 | 6.6 | 1029 | 3 | US-09-829-769-7 |
| 12 | 107 | 6.6 | 1029 | 3 | US-09-866-034-22 |
| 13 | 107 | 6.6 | 1029 | 4 | US-10-033-246-22 |
| 14 | 107 | 6.6 | 1029 | 4 | US-10-033-301-22 |
| 15 | 107 | 6.6 | 1029 | 4 | US-10-033-326-22 |
| 16 | 107 | 6.6 | 1029 | 4 | US-10-033-326-22 |
| 17 | 107 | 6.6 | 1029 | 4 | US-10-033-223-22 |
| 18 | 107 | 6.6 | 1029 | 4 | US-10-033-167-22 |
| 19 | 107 | 6.6 | 1029 | 4 | US-10-033-244-22 |
| 20 | 107 | 6.6 | 1029 | 4 | US-10-033-435-22 |
| 21 | 107 | 6.6 | 1029 | 4 | US-10-032-990-22 |
| 22 | 107 | 6.6 | 1029 | 4 | US-10-032-996-22 |
| 23 | 107 | 6.6 | 1029 | 4 | US-10-033-396-22 |
| 24 | 107 | 6.6 | 1029 | 4 | US-10-439-249-22 |
| 25 | 107 | 6.6 | 1029 | 4 | US-10-692-072-22 |
| 26 | 107 | 6.6 | 1029 | 6 | US-11-021-330-7 |
| 27 | 107 | 6.6 | 1029 | 6 | US-11-021-329-7 |

| | | | | | | |
|----|------|-----|------|---|----------------------|---------------------|
| 28 | 100 | 6.2 | 552 | 3 | US-09-833-245-1335 | Sequence 1335, Ap |
| 29 | 97.5 | 6.0 | 1548 | 4 | US-10-369-493-6347 | Sequence 6347, Ap |
| 30 | 96 | 5.9 | 1462 | 4 | US-10-437-963-184121 | Sequence 184121, Ap |
| 31 | 94.5 | 5.8 | 298 | 4 | US-10-235-026-2 | Sequence 2, Appl |
| 32 | 94.5 | 5.8 | 298 | 4 | US-10-807-897-27 | Sequence 27, Appl |
| 33 | 94.5 | 5.8 | 298 | 4 | US-10-839-882-18 | Sequence 18, Appl |
| 34 | 94.5 | 5.8 | 600 | 4 | US-10-369-493-20878 | Sequence 20878, A |
| 35 | 92.5 | 5.7 | 779 | 4 | US-10-343-663A-25 | Sequence 418, Appl |
| 36 | 92.5 | 5.7 | 779 | 5 | US-10-370-715B-418 | Sequence 418, Appl |
| 37 | 92.5 | 5.7 | 779 | 5 | US-10-852-335A-165 | Sequence 165, Appl |
| 38 | 92 | 5.7 | 416 | 3 | US-09-808-387-38 | Sequence 38, Appl |
| 39 | 91.5 | 5.7 | 2548 | 3 | US-09-851-682A-1 | Sequence 1, Appl |
| 40 | 91 | 5.6 | 465 | 3 | US-09-764-864-1278 | Sequence 1278, Ap |
| 41 | 90.5 | 5.6 | 545 | 3 | US-10-369-493-19756 | Sequence 19756, A |
| 42 | 90 | 5.6 | 943 | 4 | US-10-437-963-168658 | Sequence 168658, A |
| 43 | 90 | 5.6 | 974 | 4 | US-10-369-493-21047 | Sequence 21047, A |
| 44 | 89.5 | 5.5 | 670 | 5 | US-10-631-467-863 | Sequence 863, Appl |
| 45 | 89.5 | 5.5 | 690 | 5 | US-10-631-467-862 | Sequence 862, Appl |

ALIGNMENTS

RESULT 1
US-09-889-325-4
; Sequence 4, Application US/09889325
; Publication No. US20040224312A1
; GENERAL INFORMATION:
; APPLICANT: AGES Research Institute, Co., Ltd.
; TITLE OF INVENTION: Gene causative of Rothmund-Thomson syndrome
; TITLE OF INVENTION: and its gene product
; FILE REFERENCE: A1-003PCT
; CURRENT APPLICATION NUMBER: US/09/889,325
; PRIOR APPLICATION NUMBER: JP 1999-11218
; PRIOR FILING DATE: 1999-01-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-889-325-4

| Query Match | 100.0% | Score 1617 | DB 3 | Length 1208 |
|-----------------------|--------|---|------|--------------|
| Best Local Similarity | 100.0% | Pred. No. 4.4e-151 | | |
| Matches | 302 | Conservative | 0 | Mismatches 0 |
| | | Indels | 0 | Gaps 0 |
| QY | 1 | QLTVALDMPERAEITLLCYELHPHMLLELTATYTCRLNCPGPAQLQALAHRCPEL | 60 | |
| DB | 907 | QLTVALDMPERAEITLLCYELHPHMLLELTATYTCRLNCPGPAQLQALAHRCPEL | 966 | |
| QY | 61 | AVCLAAQQLPEDPGQSSSVFEDMTLVDSMGWELASVRRALCOLQDWHDEPRTGRTGV | 120 | |
| DB | 967 | AVCLAAQQLPEDPGQSSSVFEDMTLVDSMGWELASVRRALCOLQDWHDEPRTGRTGV | 1026 | |
| QY | 121 | LVSESELAFHLSPPDILTAEKDOI CDPLYGVAORBOALRLRTQAFHVAFSPSCG | 180 | |
| DB | 1027 | LVSESELAFHLSPPDILTAEKDOI CDPLYGVAORBOALRLRTQAFHVAFSPSCG | 1086 | |
| QY | 181 | PCLEQDDESRSTRLDLGRYFESEGOEPGGMEDAQGPQCARLODMEDQVRCDIROF | 240 | |
| DB | 1087 | PCLEQDDESRSTRLDLGRYFESEGOEPGGMEDAQGPQCARLODMEDQVRCDIROF | 1146 | |
| QY | 241 | LSLRPEEKSSAAVARIFFIGISPCYPAQVYQDDRFRKYLHLSFHLVGLATELLQV | 300 | |
| DB | 1147 | LSLRPEEKSSAAVARIFFIGISPCYPAQVYQDDRFRKYLHLSFHLVGLATELLQV | 1206 | |
| QY | 301 | AR 302 | | |
| DB | 1207 | AR 1208 | | |

```
RESULT 2
US-10-723-860-4382
; Sequence 4382, Application US/10723860
; Publication No. US20040253606a1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882, 0193, NPUS01
; CURRENT APPLICATION NUMBER: US/10/723, 860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429, 739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4382
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4382

Query Match      100.0%; Score 1617; DB 5; Length 1208;
Best Local Similarity 100.0%; Pred. No. 4,4e-151;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLTVALDMPBEALETLLCYLELHPHMLBLATTTTHCRINCPCGPAQLALAHRCPL 60
Db 907 QLTVALDMPBEALETLLCYLELHPHMLBLATTTTHCRINCPCGPAQLALAHRCPL 966

QY 61 AVCLAQQLPDPGGSSSVFEDWVKLVDSMGWELASVRALCOLQMDHEPRTGVRGTV 120
Db 967 AVCLAQQLPDPGGSSSVFEDWVKLVDSMGWELASVRALCOLQMDHEPRTGVRGTV 1026

QY 121 LVFSELAFLHRSBGDLTAEEKDOI CDPLYGRVQAREQALRLRTFOAFHSAVPSCG 180
Db 1027 LVFSELAFLHRSBGDLTAEEKDOI CDPLYGRVQAREQALRLRTFOAFHSAVPSCG 1086

QY 181 PCLEQODEERSTRKDLIGRYFEEBEGQEPGMEADAGPEPGQARLDWEDOVRCIDIRQF 240
Db 1087 PCLEQODEERSTRKDLIGRYFEEBEGQEPGMEADAGPEPGQARLDWEDOVRCIDIRQF 1146

QY 241 LSLRPEKFSRAVARIFHGIGSPCYPAQYGGDRRFWRKYLHLSPFALVGLATEBELLQV 300
Db 1147 LSLRPEKFSRAVARIFHGIGSPCYPAQYGGDRRFWRKYLHLSPFALVGLATEBELLQV 1206

QY 301 AR 302
Db 1207 AR 1208

; ORGANISM: Homo sapiens
RESULT 3
US-10-882-405-4
; Sequence 4, Application US/10882405
; Publication No. US20040259147a1
; GENERAL INFORMATION:
; APPLICANT: AGENE Research Institute, Co., Ltd.
; TITLE OF INVENTION: Gene causative of Rothmund-Thomson syndrome
; FILE REFERENCE: AI-003PCT
; CURRENT APPLICATION NUMBER: US/10/882, 405
; PRIOR FILING DATE: 2004-07-02
; PRIOR APPLICATION NUMBER: US/09/889, 325
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: JP 1999-11218
; PRIOR FILING DATE: 1999-01-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-10-882-405-4

Query Match      100.0%; Score 1617; DB 5; Length 1208;
Best Local Similarity 100.0%; Pred. No. 4,4e-151;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLTVALDMPBEALETLLCYLELHPHMLBLATTTTHCRINCPCGPAQLALAHRCPL 60
Db 907 QLTVALDMPBEALETLLCYLELHPHMLBLATTTTHCRINCPCGPAQLALAHRCPL 966

QY 61 AVCLAQQLPDPGGSSSVFEDWVKLVDSMGWELASVRALCOLQMDHEPRTGVRGTV 120
Db 967 AVCLAQQLPDPGGSSSVFEDWVKLVDSMGWELASVRALCOLQMDHEPRTGVRGTV 1026

QY 121 LVFSELAFLHRSBGDLTAEEKDOI CDPLYGRVQAREQALRLRTFOAFHSAVPSCG 180
Db 1027 LVFSELAFLHRSBGDLTAEEKDOI CDPLYGRVQAREQALRLRTFOAFHSAVPSCG 1086

QY 181 PCLEQODEERSTRKDLIGRYFEEBEGQEPGMEADAGPEPGQARLDWEDOVRCIDIRQF 240
Db 1087 PCLEQODEERSTRKDLIGRYFEEBEGQEPGMEADAGPEPGQARLDWEDOVRCIDIRQF 1146

QY 241 LSLRPEKFSRAVARIFHGIGSPCYPAQYGGDRRFWRKYLHLSPFALVGLATEBELLQV 300
Db 1147 LSLRPEKFSRAVARIFHGIGSPCYPAQYGGDRRFWRKYLHLSPFALVGLATEBELLQV 1206

QY 301 AR 302
Db 1207 AR 1208

; ORGANISM: Homo sapiens
RESULT 4
US-10-756-149-5768
; Sequence 5768, Application US/10756149
; Publication No. US20050181375a1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756, 149
; PRIOR FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5768
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-756-149-5768

Query Match      100.0%; Score 1617; DB 5; Length 1208;
Best Local Similarity 100.0%; Pred. No. 4,4e-151;
Matches 302; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLTVALDMPBEALETLLCYLELHPHMLBLATTTTHCRINCPCGPAQLALAHRCPL 60
Db 907 QLTVALDMPBEALETLLCYLELHPHMLBLATTTTHCRINCPCGPAQLALAHRCPL 966

QY 61 AVCLAQQLPDPGGSSSVFEDWVKLVDSMGWELASVRALCOLQMDHEPRTGVRGTV 120
Db 967 AVCLAQQLPDPGGSSSVFEDWVKLVDSMGWELASVRALCOLQMDHEPRTGVRGTV 1026

QY 121 LVFSELAFLHRSBGDLTAEEKDOI CDPLYGRVQAREQALRLRTFOAFHSAVPSCG 180
Db 1027 LVFSELAFLHRSBGDLTAEEKDOI CDPLYGRVQAREQALRLRTFOAFHSAVPSCG 1086

QY 181 PCLEQODEERSTRKDLIGRYFEEBEGQEPGMEADAGPEPGQARLDWEDOVRCIDIRQF 240
Db 1087 PCLEQODEERSTRKDLIGRYFEEBEGQEPGMEADAGPEPGQARLDWEDOVRCIDIRQF 1146

QY 241 LSLRPEKFSRAVARIFHGIGSPCYPAQYGGDRRFWRKYLHLSPFALVGLATEBELLQV 300
Db 1147 LSLRPEKFSRAVARIFHGIGSPCYPAQYGGDRRFWRKYLHLSPFALVGLATEBELLQV 1206
```

| | | | |
|----|------|--|------|
| Db | 1147 | LSLRPEKFSRAVARIFHGIGSPCYAQYAGDPRFRFKYLIHLSFHALVGLATELLQV | 1206 |
| Qy | 301 | AR | 302 |
| | 1207 | AR | 1208 |
| Db | | | |

RESULT 5
US-11-084-955-2
; Sequence 2, Application US/11084955
; Publication No. US20050183149A1

```

:
:
: APPLICANT: National Institute of Radiological Sciences
: TITLE OF INVENTION: Targeted knockout mouse for human Rothmund-Thomson syndrome and
: TITLE OF INVENTION: preparation method thereof
: FILE REFERENCE: YIL-0610
: CURRENT APPLICATION NUMBER: US/11/084,955
: CURRENT FILING DATE: 2005-03-21
: PRIOR APPLICATION NUMBER: JP 2003-185409
: PRIOR FILING DATE: 2003-06-27
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 2
: LENGTH: 1216
: TYPE: prt
: ORGANISM: Mus musculus
: US-11-084-955-2

```

| | | | | |
|-----------------------|------------------|---------------------|-----------|--------------|
| Query Match | 70.8%; | Score 1144.5; | DB 6; | Length 1216; |
| Best Local Similarity | 73.2%; | Pred. No. 4.1e-104; | | |
| Matches 218; | Conservative 26; | Mismatches 53; | Indels 1; | Gaps 1; |

| | | | |
|----|------|--|------|
| QY | 1 | QYVQALDMPBEALFETLLCYLEHPHMLLELATYTCRLNCPGSPNOALAHRCPL | 60 |
| Db | 916 | OSTVQALDMPTEALFETLLCYLEHPHMLLELPPTYAQCHLHCGSGAQALAHRCPL | 975 |
| QY | 61 | AVCIAQQLPEDPGQSSSVFEDMWKLVDSMGWELASVRALCOLQMDHEPTGYRRGTGV | 120 |
| Db | 976 | AACAKMPPKDTSQGRSLSEFGVEVLDSMGKCLASVQAOLHOLKMPDEPKKGAQOGTGV | 1035 |
| QY | 121 | LVPSFELAFHLRSPGDLFAEEKDOIQDPLYGRVQARERQALRLRRTQAFLSVAFPSCG | 180 |
| Db | 1036 | LVKSEELAFHLHSRGLDLDEEKDOIQDPLYNRVQAREKALAHHQMSKARSVAFPSCG | 1095 |
| QY | 181 | PCLQGQDEERSTRLKDLIGRYFEEBEGQEPGAMEDAQSPFGQALQDMDQVRCDIROF | 240 |
| Db | 1096 | PCLQGSNEHNSQVTLVSYYFEEBEE-ABEETMTDTPQPKDQIQLQDMDQIRDRVAQL | 1154 |
| QY | 241 | LSLRPEEKSSRAVARIFHGIGSPCYPAQOYGGODRFRKYLHLHSFHLVGLATEELL | 298 |
| Db | 1155 | LSLRPEEKSSGRAVARIFHGIGLSPCYPAQOYGLDRFRKYLHLDFHLHMLATEELL | 1212 |

```

1  RESULT 6
2  US-11-097-143-16743
3  ; Sequence 16743, Application US/11097143
4  ; Publication No. US20050208558A1
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Venter, J. Craig
7  ; APPLICANT: et al.
8  ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
9  ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
10 ; TITLE OF INVENTION: DROSOPHILA GENES.
11 ; FILE REFERENCE: C1000728
12 ; CURRENT APPLICATION NUMBER: US/11/097,143
13 ; CURRENT FILING DATE: 2005-04-04
14 ; PRIOR APPLICATION NUMBER: 60/157,832
15 ; PRIOR FILING DATE: 1999-10-05
16 ; PRIOR APPLICATION NUMBER: 60/160,191
17 ; PRIOR FILING DATE: 1999-10-19
18 ; PRIOR APPLICATION NUMBER: 60/161,932
19 ; PRIOR FILING DATE: 1999-10-28
20 ; PRIOR APPLICATION NUMBER: 60/164,769

```

```

? PRIOR FILLING DATE: 1999-11-12
? PRIOR APPLICATION NUMBER: 60/173,383
? PRIOR FILING DATE: 1999-12-28
? PRIOR APPLICATION NUMBER: 60/175,693
? PRIOR FILING DATE: 2000-01-12
? PRIOR APPLICATION NUMBER: 60/184,831
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: 60/191,637
? PRIOR FILING DATE: 2000-03-23
? NUMBER OF SEQ ID NOS: 43008
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 16743
? LENGTH: 1579
? TYPE: PRT
? ORGANISM: DROSOPHILA
US-11-097-143-16743

```

| | | | | |
|-----------------------|-------|--------------------|-------|-----------------|
| Query Match | 26.4% | Score 427; | DB 6; | Length 1579; |
| Best Local Similarity | 31.1% | Pred. No. 1.2e-32; | | |
| Matches | 95; | Conservative | 62; | Mismatches 122; |
| | | | | Indels 26; |
| | | | | Gaps 6; |

| | | | |
|----|------|--|------|
| Qy | 3 | TVQALNDPBEKIEITLLCTSLNHPNMYELATTYTHCRINCPGPAQIOALAHNCSPRLAV | 62 |
| Db | 1290 | TVEMLDTPAENISTILLCTMELDPWICISVSSAYVMAKVISYGGPKYLKNAKCKCPRLAM | 1349 |
| Qy | 63 | CLAQGLPEDP-GGGSSSVEFDMVVLVDMSGWELASVRRALCOLQMDHEPRYGVRTGTVL | 121 |
| Db | 1350 | ALAQIQDKTKFKEDNSNIIEFSDVIDIAAGIGNSGVVKYQKLDLEM--VKVNGYPRSPIT | 1407 |
| Qy | 122 | VEFSELAPHLRSPGDLTAEEKQOICDPLYGVQAREQOALRLRRTQAFHSVAPPSGCP | 181 |
| Db | 1408 | VSFYDVGPRIKVPDGFESIEDNALDLYTRSVKQERTQOLIQLQVYAHGLAAVAVSSCGQ | 1467 |
| Qy | 182 | CLEEQ-DEBKSTRKLDLGRYFEESBQ-----EPGEMEDAQBPBQALQMDQDVR | 234 |
| Db | 1468 | CCNADFPQDREGEOLKAVRNYPANDYQDULELETPSNVPD-----ENI I | 1512 |
| Qy | 235 | CDIQQFSLRBEKFEKSSRAVARIRFHGSPCCYPAQVYQODBRFMRKYHLSEFHALVGLAT | 294 |
| Db | 1513 | DDVALANIMYDNTFTGTGNARIRFHGIMSPNYPAVIWGR-CKFWRAHVKVDVFNRLHLAN | 1571 |
| Qy | 295 | BELLQ 299 | |
| Db | 1572 | MATIK 1576 | |

```

: RESULT 7
: US-10-437-963-121617
: Sequence 121617, Application US/10437963
: Publication NO. US20040123343A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: APPLICANT: Wu, Wei
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Barbazuk, Brad
: APPLICANT: Li, Ping
: TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53221)B
: CURRENT APPLICATION NUMBER: US/10/437,963
: CURRENT FILING DATE: 2003-05-14
: NUMBER OF SEQ ID NOS: 204966
: SEQ ID NO 121617
: LENGTH: 874
: TYPE: PR1
: ORGANISM: Oryza sativa
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(874)
: OTHER INFORMATION: unsure at all Xaa locations

```

```
FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24624C.1.pep
US-10-437-963-121617

Query Match      10.5%; Score 169; DB 4; Length 874;
Best Local Similarity 20.9%; Pred. No. 2.4e-07;
Matches 68; Conservative 58; Mismatches 118; Indels 82; Gaps 12;

QY 1 QLTVALDMPBEAETLLCTYLHPHMLLATTYHGRNLCPGPAOLALHRCPL 60
DB 598 ELTSRKEDIKKEEVLLITLTQLEIDQOYIRLLPQSVYCTI-----YFHKIS PQ 646
QY 61 AVC-----IAOQLPBDPGGSSSVFEDWYKLVDSMGWELASVRBALCQIOWDHERPTGRR 116
DB 647 LLDADKDLIRSVLNRSEMKDGHYFDIPRIANDLITINNEY-----FDHLHK----- 693
QY 117 GTGLVLEFS-ELAFHLRSPG-----DLTAEEKQICDPLYGRVQAEERQALRL 165
DB 694 -----LKFSGEISPELKDPAVCYIIMRPDPFNLSAN-----LTKWLSVESSKISKLD 743
QY 166 RTFQAFHSVAFPSG-----GPCLEQODEERSTRLDLGRYFEEBEGEPGME 214
DB 744 AMF-ALANFAVKGCKRTGCGSGSQTHTPCIQKIME-----YFSKDGTSENDC- 790
QY 215 DAQGPBEGQARLQDMEDQVRCIDIRQFLSLRPEKSSRAVARIFHIGSPCYPAQVYQD 274
DB 791 -----RTQLQKSSPPLQMDIVKFIOGNSPAKFTPRAVARIMHGISSPAFVSATWSKN 842
QY 275 RRFWRKYLHLSPHALVGLATEELLQV 300
DB 843 -HFWGRYVEVDPLVMEAKAEIVKL 867

RESULT 8
US-10-425-114-68934
; Sequence 68934, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yinhua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68934
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73192E07_FIL.pep
US-10-425-114-68934

Query Match      9.0%; Score 145.5; DB 4; Length 206;
Best Local Similarity 25.0%; Pred. No. 7.6e-06;
Matches 50; Conservative 32; Mismatches 69; Indels 49; Gaps 9;

QY 123 EFS-ELAFHLRSPG-----DLTAEEKQICDPLYGRVQAEERQALRLRRTQAF 171
DB 29 QFSGEISFELKDPAICIVILKKPDDLNALAD-----ITRWLSVEVSKISKLDAMPD-L 82
QY 172 HSAVAFPSG-----GPCLEQODEERSTRLDLGRYFEEBEGEPGMEDAQPE 220
DB 83 ANFAVKGCKRTDGGSGSQTHTPCIQK-----IIQVFSKNYSTSDSG--QCTOPQ 129
QY 221 PGQARLQDMEDQVRCIDIRQFLSLRPEKSSRAVARIFHIGSPCYPAQVYQD RFRMRK 280
DB 130 RGSFPLO-----ADIKVFLKNSPAKFTPRAVARIMHGISSPAFVSATWSKN-HFWGR 181
```

```
QY 281 YLHLSFHALVGLATEELLQV 300
DB 182 YLEVDFFVWMEAKAEIVKL 201

RESULT 9
US-10-425-115-338629
; Sequence 338629, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53122)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 338629
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_71997C.1.pep
US-10-425-115-338629

Query Match      9.0%; Score 145.5; DB 4; Length 215;
Best Local Similarity 25.0%; Pred. No. 8e-06;
Matches 50; Conservative 32; Mismatches 69; Indels 49; Gaps 9;

QY 123 EFS-ELAFHLRSPG-----DLTAEEKQICDPLYGRVQAEERQALRLRRTQAF 171
DB 38 QFSGEISFELKDPAICIVILKKPDDLNALAD-----ITRWLSVEVSKISKLDAMPD-L 91
QY 172 HSAVAFPSG-----GPCLEQODEERSTRLDLGRYFEEBEGEPGMEDAQPE 220
DB 92 ANFAVKGCKRTDGGSGSQTHTPCIQK-----IIQVFSKNYSTSDSG--QCTOPQ 138
QY 221 PGQARLQDMEDQVRCIDIRQFLSLRPEKSSRAVARIFHIGSPCYPAQVYQD RFRMRK 280
DB 139 RGSFPLO-----ADIKVFLKNSPAKFTPRAVARIMHGISSPAFVSATWSKN-HFWGR 190
QY 281 YLHLSFHALVGLATEELLQV 300
DB 191 YLEVDFFVWMEAKAEIVKL 210

RESULT 10
US-10-476-924-22
; Sequence 22, Application US/10476924
; Publication No. US20040152093A1
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry; Ding, Li;
; APPLICANT: Baughn, Mariah R.; Lal, Preeti G.;
; APPLICANT: Yue, Huidun; HAPALA, April J.A.;
; APPLICANT: Lee, Ernestine A.; ISON, Craig H.;
; APPLICANT: BECHA, Shanya D.; GURUBAUN, Rajagopal;
; APPLICANT: EMBELING, Brooke M.; GRIFFIN, Jennifer A.;
; APPLICANT: TANG, Y. Tom; Luo, Dying Ana M.;
; APPLICANT: YAO, Monique G.; CHAWLA, Nandinder K.;
; APPLICANT: RAMKUMAR, Jayalakshmi; GANDHI, Ameena R.;
; APPLICANT: Lee, Soo Yeun; RICHARDSON, Thomas W.;
; APPLICANT: YANG, Junning; ELIOTT, Vicki S.;
; APPLICANT: LU, Yan; THANGAVELU, Kavitha;
; APPLICANT: HE, Ann; AZIMZAI, Yalda;
; APPLICANT: RAUMANN, Brigitte E.; SWARNAKAR, Anita;
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0960 USN
; CURRENT APPLICATION NUMBER: US/10/476,924
```

```

; CURRENT FILING DATE: 2003-11-04
; PRIOR APPLICATION NUMBER: PCT/US02/14276
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US 60/288,598
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/291,776
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/292,172
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/293,564
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1532441CD1
US-10-476-924-22

```

```

Query Match 6.6%; Score 107; DB 4; Length 751;
Best Local Similarity 27.0%; Pred. No. 0.29;
Matches 70; Conservative 23; Mismatches 106; Indels 60; Gaps 10;

```

```

QY 11 BEATITLCYELAHMHMLLATTYTT---HCLNCPGPGPAQLOALAHRCPIAVCLA 65
DB 159 BEARERQAQCPQHGEALRFQPCPSQLCRECLD-----PHLDHCLPIAEAVR 210
QY 66 QQLPEDPGQSSSVFEDWKLVDSMGWELASVR---ALCOLMD-----HEPRTG 113
DB 211 ARPGLGEG-----LAGVDNNLVELAEARVEKEALAEARQAVGTQVEEAEG 261
QY 114 VARGTGVVEFSELAFLHLSRPGDLTAEBKQICDPLYGVRQ-AREQALAR----- 163
DB 262 VLR--ALLAQKQEVGQLRAHVEAEBAARERLAELEGEGVARRAAAARAVLSIGREA 319
QY 164 -----LRTFOAFHSVAFPSGCPCLSEQDERSTLK---DLGTYFEFEEQEG 211
DB 320 EILSLEGAIAQRIRQIOGCPMAFPAPCLLPQLLEHPLGLDKXCHLLRLSFEQGPQKXG 379
QY 212 GMEGA--OGPEPGOARLOD 228
DB 380 GKQAGTGGESQSERED 398

```

```

RESULT 11
US-09-929-769-7
; Sequence 7, Application US/09929769
; Publication No. US20030055224A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Polakis, Paul
; APPLICANT: Shou, Jianyong
; APPLICANT: Smith, Victoria
; APPLICANT: Soriano, Robert
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITL OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: F5007R1-US
; CURRENT APPLICATION NUMBER: US/09/929,769
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/089,653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/090,355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/104,257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10

```

```

; PRIOR APPLICATION NUMBER: 60/141,037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/20111
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: PCT/US00/00376
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PCT/US00/04342
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/13705
; PRIOR FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/20118
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/888,257
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-929-769-7

```

```

Query Match 6.6%; Score 107; DB 3; Length 1029;
Best Local Similarity 26.2%; Pred. No. 0.44;
Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

```

```

QY 29 LEIATTTYHCRINCPGPAQLOALAHRCPIAVCLAQUL-----PEDPGQGS--SSV 79
DB 245 LQALATL-----LSSPHGALVMSH-HRSHFLACPLRLQCOYQRCVQDGTGFSLSFLKV 298
QY 80 EFDVVKLVDSMGWELASVRALCOL--QMDHEPRTGVRGTGVVEFSELAH----- 130
DB 299 LLQMLQWIDSPGVGEGPILRAQLBMLASQASAGRRSLDVRG--GLRLABALAFRODLEVS 357
QY 131 -----LNSPGDLTAEBKQICDPLYGVRQVAREQALRLRTFOAFHSVA--FPSC 179
DB 358 STVRVIAITLRS-GRQCSVEPDLISKVLQGLEIVASPHLEILLTAFSATADASPPFAC 416
QY 180 GPCLEQDERSTRLKDLIGRYFEER---GOEP--GMEGAQPEPGOARLODMEQV 233
DB 417 KPVVV-----VSSLL--LQEEEPDAGKPADGSGLEAVVLGSSGLLVMLENL 464
QY 234 RCDIROFLSLRPEEK-----FSSRAVARIFHGIDPCYPAQVYGDRRFRKRY----- 281
DB 465 DEPV---VSSCPDQLRLILFSRR-----KKGQAQVPS-----FRPYLLTLFTHQ 506
QY 282 -----LHLSFHALVGLATEE 296
DB 507 SSWPTLHQCIRVLLGKSRQ 526

```

```

RESULT 12
US-09-866-034-22
; Sequence 22, Application US/09866034
; Publication No. US20030170864A1
; GENERAL INFORMATION:

```

APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Pong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1C1
CURRENT APPLICATION NUMBER: US/09/866,034
CURRENT FILING DATE: 2001-05-25
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 22
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-09-866-034-22

Query Match 6.6%; Score 107; DB 3; Length 1029;
Best Local Similarity 26.2%; Pred. No. 0.44;

Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

QY 29 LELIATYYTHCRINCPGPAQLOALAHRCPLAVCLAQOL-----PEDPGQS--SSV 79
DB 245 LQALATL-----LSPHFGALVWSM-HRSHFLACPLRLQCYORCVPDGTGSSSLFLKV 298
QY 80 EFDWVKLVDSKMGWELASVRALCOL--QMDHEPRTGVRRGTGVIVSESLAFH----- 130
DB 299 LQWLQWLDSPPGVGGPLRAQLRMLASQASAGRRLSDVRG--GLRLALEALAFRODLEVVS 357
QY 131 -----LRSPDGLAEKQDQICDFLYGRVQARERQALRLRRTQAFHSVA--PPSC 179
DB 358 STYRAVATATLRS--GEQCSVEPDLISKVLQGLIEVRSPHLEELTAFFSATADASPFPAC 416
QY 180 GPCLEQDDEERSTLXDLGRYFEEB---GQEP---GGMADAGPREGQARLDQWBEQV 233
DB 417 KPVVV-----VSSIL--LQEEPLACGKAGDGSGLEAVRLGSSGLVDWLEML 464
QY 234 RCDIRQPLSLRPERK---FSSRAVARIFFHGIGSPCYAQVYGODRRFWRKY----- 281
DB 465 DPEV---VSSCPDLQRLRLFERR-----KKGQAQVPS-----FRYLLTLFTHQ 506
QY 282 -----LHLSFHALVGLATBE 296
DB 507 SSWPTLHQICIRVLGKSRQ 526

RESULT 13

US-10-033-246-22
Sequence 22, Application US/10033246
Publication No. US20020098505A1
GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ferrara, Napoleone
APPLICANT: Pong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2930R1C12
CURRENT APPLICATION NUMBER: US/10/033,246
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: 60/095,325
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/112,851
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,145
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113,511
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/115,558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115,733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119,341
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119,537
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119,965
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: 60/162,506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170,262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/187,202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/28634
PRIOR FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: PCT/US00/14941
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 22
LENGTH: 1029
TYPE: PRT
ORGANISM: Homo sapiens
US-10-033-246-22

Query Match 6.6%; Score 107; DB 4; Length 1029;
Best Local Similarity 26.2%; Pred. No. 0.44;

Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;

QY 29 LELIATYYTHCRINCPGPAQLOALAHRCPLAVCLAQOL-----PEDPGQS--SSV 79
DB 245 LQALATL-----LSPHFGALVWSM-HRSHFLACPLRLQCYORCVPDGTGSSSLFLKV 298
QY 80 EFDWVKLVDSKMGWELASVRALCOL--QMDHEPRTGVRRGTGVIVSESLAFH----- 130
DB 299 LQWLQWLDSPPGVGGPLRAQLRMLASQASAGRRLSDVRG--GLRLALEALAFRODLEVVS 357
QY 131 -----LRSPDGLAEKQDQICDFLYGRVQARERQALRLRRTQAFHSVA--PPSC 179
DB 358 STYRAVATATLRS--GEQCSVEPDLISKVLQGLIEVRSPHLEELTAFFSATADASPFPAC 416

```
QY 180 GPCLEQDQDERSTRKDLIGRYFEER---GOEP---GEMDAQGPFGQARLODMEDQV 233
Db 417 KPVVV-----VSSLL---LQEEPLAGKPGADGGSLEAVNLGSSGLVDMLEML 464
QY 234 RCDIRQFLSLRPEEK---FSSRAVARIFHGIGSPCYPAQVYGODRRFRKX----- 281
Db 465 DPEV---VSSCPDQLRLLFSSR-----KKGQAQVPS-----FRPYLLTLFTHQ 506
QY 282 ----LHLSFHALVGLATEE 296
Db 507 SSWPTLHQCIRVLGKSREQ 526
```

```
RESULT 14
US-10-033-301-22
; Sequence 22, Application US/10033301
; Publication No. US20020098506A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2930R1C6
; CURRENT APPLICATION NUMBER: US/10/033,301
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
```

```
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 22
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-301-22
```

```
Query Match 6.6%; Score 107, DB 4, Length 1029,
Best Local Similarity 26.2%; Pred. No. 0.44,
Matches 84; Conservative 34; Mismatches 112; Indels 90; Gaps 18;
```

```
QY 29 LELATTYTHGRNLCPGSPADQLAHRCPLAVCLAQQL-----PEDGQGS--SSV 79
Db 245 LQALATL-----LSPFHGALVMSW-HRSHPACPLRLQLOQYORCVPDGTFSSSLFKV 298
QY 80 EPDVKLVDSMGWELASVRALCOL--QWDHEPRTGVRGTVLVEFSELAFH----- 130
Db 299 LIQMLQWIDSPGVGEPFLRAQLRMILASQASGRRLSDVRG-GLRLAEALAFRODLEVVS 357
QY 131 -----LRSFGDLTAEEKDQICDPLYGAVQAREQALRLRRTFOAFHSVA--FPSC 179
Db 358 STVRAVINTKRS-GRQCVPEPDLISKVLQGLIEVNSPHLEILTFASFATADAASPFPAC 416
QY 180 GPCLEQDQDERSTRKDLIGRYFEER---GOEP---GEMDAQGPFGQARLODMEDQV 233
Db 417 KPVVV-----VSSLL---LQEEPLAGKPGADGGSLEAVNLGSSGLVDMLEML 464
QY 234 RCDIRQFLSLRPEEK---FSSRAVARIFHGIGSPCYPAQVYGODRRFRKX----- 281
Db 465 DPEV---VSSCPDQLRLLFSSR-----KKGQAQVPS-----FRPYLLTLFTHQ 506
QY 282 ----LHLSFHALVGLATEE 296
Db 507 SSWPTLHQCIRVLGKSREQ 526
```

```
RESULT 15
US-10-033-326-22
; Sequence 22, Application US/10033326
; Publication No. US20020098507A1
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2930R1C8
; CURRENT APPLICATION NUMBER: US/10/033,326
; PRIOR FILING DATE: 2001-12-27
; PRIOR APPLICATION NUMBER: 60/095,325
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/112,851
; PRIOR FILING DATE: 1998-12-16
```

```
; PRIOR APPLICATION NUMBER: 60/113,145
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113,511
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/115,558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115,733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119,341
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119,537
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119,965
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 60/162,506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170,262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/187,202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/28634
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: PCT/US00/14941
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 22
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-033-326-22
```

Query Match 6.6%; Score 107; DB 4; Length 1029;

Best Local Similarity 26.2%; Pred. NO. 0.44; Mismatches 112; Indels 90; Gaps 18;

```
29 LELIATYTHGRLNCPGGPAQLALHRCPLAVLAQOL-----PEDPGQS--SSV 79
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 245 LQALATL-----LSPHGALVMSM-HRSHLACPLRLQLCYQRCVQDTPGSSPLKV 298
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 80 EFDNVKLVDSKMGELASVRLCOL--QWDHEPRTGVNRGTGLVEFSELAFH----- 130
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 299 LQWLQWLDSFVGGVGLFAQLRLMLASQASAGRRLSDVRG--GLRLMALAFRODLEVVS 357
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 131 -----LRSPGDLTAEEKDOICDPLVGRVQAREKQALRLRTFQAFHVA--PPSC 179
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 358 STYRAVAATLKS--GEQSVEPDLISKVLQGLIEVRSPLHEELLTAFFSATDAASPPFAC 416
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 180 GPCLEQODDERSTRKDLGRYFEER---GQEP---GAMEDAQBPBGQARLDQWEDQV 233
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 417 KPVVV-----VSSLL---LQEEPLAGKGAGDGSLEAVRLGPSGLVDWLEML 464
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 234 RCDIRQFLSLRPEEK---FSSRAVARIFFHGIGSPCYPAQVYGQDRRFWRKY----- 281
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 465 DPEV---VSSCPDQLRLRLFRR-----KKGQAQVPS-----FRPYLLTLFTHQ 506
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

QY 282 -----LHLSFHAIVGLATEE 296
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 507 SSMPTHQICIRVLGKSREQ 526

Search completed: December 27, 2005, 22:12:35
Job time : 45.6 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: December 27, 2005, 21:54:12 / Search time 3 Seconds
(without alignments)
718.036 Million cell updates/sec

Title: US-09-889-325-4_COPY_907_1208
Perfect score: 1617
Sequence: 1 QLVQALDMPBAITLLCY.....HLSPHATVGLATEELLQVAR 302

Scoring table: BLOSSUM62
Gapop 10.0 , Gapext 0.5

Searched: 54001 seqs, 7132810 residues

Total number of hits satisfying chosen parameters: 54001

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US12_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 88.5 | 5.5 | 3674 | 7 US-11-000-463-454 | Sequence 454, App |
| 2 | 88 | 5.4 | 3375 | 7 US-11-044-111-23 | Sequence 23, App1 |
| 3 | 80.5 | 5.0 | 599 | 7 US-11-109-157A-3 | Sequence 1, App1 |
| 4 | 80.5 | 5.0 | 1686 | 7 US-11-109-157A-1 | Sequence 8056, Ap |
| 5 | 79.5 | 4.9 | 717 | 6 US-10-467-657-8056 | Sequence 940, App |
| 6 | 78 | 4.8 | 615 | 6 US-10-995-561-940 | Sequence 942, App |
| 7 | 76 | 4.7 | 562 | 7 US-11-065-943-24 | Sequence 24, App1 |
| 8 | 75.5 | 4.7 | 662 | 6 US-10-995-561-943 | Sequence 942, App |
| 9 | 75.5 | 4.7 | 702 | 6 US-10-995-561-942 | Sequence 941, App |
| 10 | 75.5 | 4.7 | 754 | 6 US-10-995-561-941 | Sequence 897, App |
| 11 | 75 | 4.6 | 366 | 7 US-11-000-463-897 | Sequence 425, App |
| 12 | 75 | 4.6 | 400 | 7 US-11-000-463-425 | Sequence 136, App |
| 13 | 74 | 4.6 | 278 | 6 US-10-131-826A-136 | Sequence 1658, Ap |
| 14 | 74 | 4.6 | 777 | 6 US-10-821-234-1658 | Sequence 773, App |
| 15 | 74 | 4.6 | 3803 | 6 US-10-995-561-773 | Sequence 771, App |
| 16 | 74 | 4.6 | 3960 | 6 US-10-995-561-771 | Sequence 774, App |
| 17 | 74 | 4.6 | 5335 | 6 US-10-995-561-777 | Sequence 779, App |
| 18 | 74 | 4.6 | 5415 | 6 US-10-995-561-774 | Sequence 775, App |
| 19 | 74 | 4.6 | 5415 | 6 US-10-995-561-775 | Sequence 776, App |
| 20 | 74 | 4.6 | 5935 | 6 US-10-995-561-776 | Sequence 65, App1 |
| 21 | 74 | 4.5 | 5935 | 7 US-11-065-943-65 | Sequence 128, App1 |
| 22 | 73.5 | 4.5 | 562 | 7 US-11-065-943-28 | Sequence 32, App1 |
| 23 | 73.5 | 4.5 | 562 | 7 US-11-065-943-32 | Sequence 12, App1 |
| 24 | 73.5 | 4.5 | 677 | 6 US-10-982-545-12 | |
| 25 | 73.5 | 4.5 | | | |

| | | | | | |
|----|------|-----|------|----------------------|-------------------|
| 26 | 73 | 4.5 | 798 | 7 US-11-107-028-3 | Sequence 3, App1 |
| 27 | 73 | 4.5 | 1068 | 6 US-10-467-657-2904 | Sequence 2904, Ap |
| 28 | 72 | 4.5 | 545 | 7 US-11-065-943-67 | Sequence 67, App1 |
| 29 | 72 | 4.5 | 562 | 7 US-11-065-943-34 | Sequence 34, App1 |
| 30 | 71.5 | 4.4 | 596 | 7 US-11-082-389-420 | Sequence 420, App |
| 31 | 71 | 4.4 | 286 | 7 US-11-063-343-22 | Sequence 22, App1 |
| 32 | 71 | 4.4 | 554 | 6 US-10-850-816-2 | Sequence 2, App1 |
| 33 | 71 | 4.4 | 562 | 7 US-11-065-943-36 | Sequence 26, App1 |
| 34 | 71 | 4.4 | 562 | 7 US-11-065-943-36 | Sequence 36, App1 |
| 35 | 71 | 4.4 | 616 | 6 US-10-982-545-5 | Sequence 5, App1 |
| 36 | 71 | 4.4 | 832 | 7 US-11-065-943-100 | Sequence 100, App |
| 37 | 71 | 4.4 | 832 | 7 US-11-007-792A-11 | Sequence 11, App1 |
| 38 | 71 | 4.4 | 832 | 7 US-11-007-642B-11 | Sequence 11, App1 |
| 39 | 70.5 | 4.4 | 562 | 7 US-11-065-943-22 | Sequence 22, App1 |
| 40 | 70.5 | 4.4 | 1388 | 6 US-10-821-234-1143 | Sequence 1143, Ap |
| 41 | 70 | 4.3 | 545 | 7 US-11-065-943-97 | Sequence 97, App1 |
| 42 | 70 | 4.3 | 691 | 6 US-10-131-826A-16 | Sequence 16, App1 |
| 43 | 70 | 4.3 | 703 | 6 US-10-821-234-963 | Sequence 963, App |
| 44 | 70 | 4.3 | 1798 | 6 US-10-995-561-1033 | Sequence 1033, Ap |
| 45 | 70 | 4.3 | 1798 | 6 US-10-995-561-1034 | Sequence 1034, Ap |

ALIGNMENTS

RESULT 1
US-11-000-463-454
; Sequence 454, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radjole T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 454
; LENGTH: 3674
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-000-463-454

Query Match 5.5%; Score 88.5; DB 7; Length 3674;
Best Local Similarity 20.6%; Pred. No. 4.1;
Matches 68; Conservative 40; Mismatches 83; Indels 139; Gaps 16;
QY 6 ALDMPBAITLLCYELHPHMLELATTYTRCRINCPGPAQALAHRCPLAVCIA 65

Db 823 ALSPGSSLR-----NPGWSE-----ASCHPGSGAMKMA----- 853
Qy 66 QQLPBDPQSSSVSEFPMVKLV-----SMGWE-----LASVRA----- 100
Db 854 --LPAEP-----DDFPDNTLITQTOHLSQDYESTLRALQALRRLAEAMALFGPCSSCG 906
Qy 101 -----LQOLQMDHEPRTGVRGTGVLYVESELAFLHRS--P 134
Db 907 ELQJLMEKQVTLQVQPOADTLEVMQJLKENF--LTLAVGGMABVSSAEQJLRQYP 965
Qy 135 GDLTAEEKDQ--ICDFLYGRVQARERQALRLRTFOAFHSAFPS-----CGPCTEQODEE 189
Db 966 GNSIQIQOQOEELSQRMQQLAKRE-----KAVQLASHSVCSFLQOCCP----- 1011
Qy 190 RSTLKLGLGYFESEEGQEPGEMDAG-----PEPQ 223
Db 1012 TOVQLRVL-----LOLEALQPGSSSEDTLHALQAKTTLVLERVYFLQSVVVKVEEPGY 1067
Qy 224 ARLODMEQVACDIROFLSLRPEKFSRA 253
Db 1068 ABOPLQGV--ETLQGLIKOVQEQVQRA 1095

RESULT 2

US-11-044-111-23
; Sequence 23, Application US/11044111
; Publication No. US20050272362A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Wen
; APPLICANT: Strasburg, Gale
; APPLICANT: Linz, John
; TITLE OF INVENTION: Genetic Test for PSE-Susceptible Turkeys
; FILE REFERENCE: MSU-09308
; CURRENT APPLICATION NUMBER: US/11/044,111
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 3375
; TYPE: PRT
; ORGANISM: Meleagris gallopavo
US-11-044-111-23

Query Match 5.4%; Score 88; DB 7; Length 3375;

Best Local Similarity 21.9%; Pred. No. 4.1; Matches 87; Conservative 35; Mismatches 126; Indels 150; Gaps 17;

Qy 4 VOALDMEBE-----AIEPLLCTLELPHHMLBLATTYHCHLNCGPGPQALQALAHRCPP 59
Db 230 MMALHIEBRKCMIMELMERQDLRFHW-----HTLKLVCALGNTRVAAHALCSHYDP 285
Qy 60 LAVCLAQQLPEDPG-----QGSSEVFPM----- 83
Db 286 SOLFAIRSPFLPRLRAGYDYLAVHLBOGVABASMSSTEFIVPMDEASKRISLPAG 345
Qy 84 -----VKLVDSMGWELASV-----RRALCOLWD 107
Db 346 GGGGVKVPGGPVGSLACLRPRPAEBCFVRPPDGRALLGPSILRALGRRALRMLR-- 403
Qy 108 HEPTG-----VRSGTVLVEFS-----ELAFHLSPPGDLTAERKQIDFLYGRV----- 153
Db 404 -EAVAGGPHARDVGGVGEVQLVPLVKLVASALAVGLARPEYAKVLMIEPRVGGGE 462
Qy 154 QARERQALRLRTFOAFHSAFPSCGPCTEQODEERSTRLKDLLGRYFESEEGQEPGGM 213
Db 463 EEEEEERRRRRKAVERA-----GBEEBEVDEEBEV-----DEEBDEE 499
Qy 214 EDAGGPEPG--QALQDMEQVACDIROF-----LSLPER--KFSRAVARIFHGIGSP 264
Db 500 EEEEGPEEGGLQMKLPSSVKLQMCNLQFCDQELQHRHVEAIVAFSEHVERLQR----- 554
Qy 265 CYPQAVYGODRRFMRKTYLHLSFHALVGLATEELLQVAV 302

Db 555 -----DQRRRYGR-----LMGAVTMSAETAR 576

RESULT 3

US-11-109-157A-3
; Sequence 3, Application US/1109157A
; Publication No. US20050277175A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: TRUNCATED ADAMTS MOLECULES
; FILE REFERENCE: 01997.030500.
; CURRENT APPLICATION NUMBER: US/11/109,157A
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/562,685
; PRIOR FILING DATE: 2004-04-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 599
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-11-109-157A-3

Query Match 5.0%; Score 80.5; DB 7; Length 599;

Best Local Similarity 23.2%; Pred. No. 2.5; Matches 82; Conservative 36; Mismatches 81; Indels 155; Gaps 20;

Qy 44 PGPAQLQALAHRCPPPLAVCAQQLPEDPG-----QGSSEVFPMVKLVDSMG----- 91
Db 2 PGSPSP-RSPAPLRLPILLLLCALAPAPAPBGRATEGRAALDIVHPRVDAGGSFLSY 60
Qy 92 --WELA-----SVRR--ALCQLQMDHEPRTGVRGTGVLYVESELAFA-----HURSP 134
Db 61 ELMPRLAKRKDVSRDAPAFYELQY-----RG-----RELPFLMTANQHLAP 104
Qy 135 G-----DLTAERKD-----QTCDFLYGRVQ----- 154
Db 105 GFVSETRRGGLGRAHTRANTPACHLGEVQDPEELBGLAIAISCDGILKGVQLSNEYF 164
Qy 155 -----AR-----ERQALRLRLRTFOAFHSAFPSCG-----PCV-----EQ 185
Db 165 IBPLDSAPARPGHAPVHVYKQAPERLA---QKGDSSASTQCVQVYVPELESRRERMEQ 221
Qy 186 QDEERSTRLKDLGRYFESEEGQEPGEMDAGPE--EQARLODMEQVACDIROFLSLR 244
Db 222 RQCMRRPRLRLHQRSVSKSKWETTLVADAKWVEYHGQFQVESY-----VLTIMMM- 273
Qy 245 PEKFSRAVARIFH--GISPCYPQAVYGODRRFMRKTYLHLSFHALVGLATEE 296
Db 274 -----VAGLFHDPSTGNP-----IHTTVRLVLEDEE 301

RESULT 4

US-11-109-157A-1
; Sequence 1, Application US/1109157A
; Publication No. US20050277175A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: TRUNCATED ADAMTS MOLECULES
; FILE REFERENCE: 01997.030500.
; CURRENT APPLICATION NUMBER: US/11/109,157A
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: 60/562,685
; PRIOR FILING DATE: 2004-04-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 1686
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-109-157A-1

Query Match 5.0%; Score 80.5; DB 7; Length 1686;

Best Local Similarity 23.2%; Pred. No. 9.1;
Matches 82; Conservative 36; Mismatches 81; Indels 155; Gaps 20.

```

QY      44  PGGPAPLQALARCPCPLAACLAQQLPEPBG-----QGSSSVFEDKWLVDSSNG----- 91
Db      2  PGGPSP-RSPAPLRLPLLLLLCALAPGAPGAPGATGGRALLDIYHPRVYDAGGSTLYS 60

QY      92  --WELA-----SVRR--ALCOLQMDHEPRYGRGTGVIVERSELAF-----HLRSP 134
Db      61  ELMPLRLRKRDVSVRDDAPAFYELOY-----RG-----RELRFNLTAQCHLLAP 104

QY      135  G-----DTAEKED-----QICDFLYGRVQ----- 154
Db      105  GFVSETRRRGGLGRAHIRAHPTRPACHLLEGVQDPELEGGLAIAISACDGLKGVFLQSNEDYF 164

QY      155  -----AR-----EPQALARLRLRTTQAFHSAVAFSCG-----PCF-----EQ 185
Db      165  IEPIDSAAPAPGAPHAQPHVYVKRAQDAERLA--QRDSSAPSPCTCGVQVYPELESRRERMEQ 221

QY      186  QDQERSTRKLQDLIGRFFEEBEGQEPGCGMEDAGSP-PPQALQLOMDQDQVYRDINQFSLR 244
Db      222  RQQRPPRLRLRLHQRSVSKEKWEVLVYAADQGVETHQPOVSEY-----VLTINNM- 273

QY      245  PEKFSRSLAVARIFH--GIGSPCYPAQVYQQRRRFMRYKYLHLSFALVGLATEE 296
Db      274  -----VAGIFHDPSIGNP-----IHTTYRVLLDDEB 301

```

RESULT 5
US-10-467-657-8056

```

1  GENERAL INFORMATION:
2  APPLICANT: CHIRON SpA
3  APPLICANT: FONTANA Maria Rita
4  APPLICANT: PIZZA Mariagrazia
5  APPLICANT: MASIGNANI Vega
6  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
7  FILE REFERENCE:
8  CURRENT APPLICATION NUMBER: US/10/467,657
9  CURRENT FILING DATE: 2003-08-11
10 PRIOR APPLICATION NUMBER: GB-0103424.8
11 PRIOR FILING DATE: 2001-02-12
12 NUMBER OF SEQ ID NOS: 9218
13 SOFTWARE: SeqwIn99, version 1.04
14 SEQ ID NO 8056
15 LENGTH: 717
16 TYPE: PR1
17 ORGANISM: Neisseria gonorrhoeae
18 US-10-467-657-8056

```

Query Match 4.9%; Score 79.5; DB 6; Length 717;

Best Local Similarity 21.5%; Pred. No. 4;
Matches 39; Conservative 28; Mismatches 69; Indels 45; Gaps 5

[illegible]

RESULT 6
US-10-995-561-940
; Sequence 940, Application US/10995561

```

1  APPLICANT: CARGILL, Michele et al.
2  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
3  TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
4  TITLE OF INVENTION: DETECTION AND USES THEREOF
5  FILE REFERENCE: CL001559
6  CURRENT APPLICATION NUMBER: US/10/995,561
7  CURRENT FILING DATE: 2004-11-24
8  NUMBER OF SEQ ID NOS: 85702
9  SOFTWARE: FastSeq for Windows Version 4.0
10 SEQ ID NO 940
11 LENGTH: 615
12 TYPE: PRT
13 ORGANISM: Homo sapiens
14 US-10-995-561-940

```

| | | | | |
|-----------------------|-----------------|---------------|-----------|------------|
| Query Match | 4.8% | Score 78 | DB 6 | Length 615 |
| Best Local Similarity | 28.1% | Pred. No. 4.6 | | |
| Matches 38 | Conservative 15 | Mismatches 58 | Indels 24 | Gaps 6 |

```

OY      22 ELAHPHMLILATTTTHCINCPCGPAPD-----QALAHRCPLAVCLAQOLPEDEG----- 73
Db      495 EVAVSWMLFFPSLSGFSLLVGVGLVGIQLARRARHLCPLPPIPCASSALHFFPGCKETW 55
OY      74 QGSSVEFPMVKLVDSMGWELASVRAL-CQLQMDHEBRTGVRRKGTGLVFESELAFLRL 13
Db      555 QMINPVPDFO-----EEASLQALVEMSMVGKRTPELKT-----ELPEGAPBLA 60
OY      133 SPGDLTAEEKDQICD 147
Db      601 LDTLESLIEDGR-CD 614

```

RESULT 7
US-11-065-943-24

```

: Sequence 24 Application US/11065943
: Publication No. US20050250131A1
: GENERAL INFORMATION:
: APPLICANT: JESTIN JEAN-LUC
: APPLICANT: VICHIER-GIERRE, SOPHIE
: APPLICANT: FERRIS, STEPHANE
: TITLE OF INVENTION: METHODS FOR OBTAINING THERMOSTABLE ENZYMES, DNA POLYMERASE I
: TITLE OF INVENTION: VARIANTS FROM THERMUS AQUATICUS HAVING NEW CATALYTIC ACTIVITIES
: TITLE OF INVENTION: METHODS FOR OBTAINING THE SAME, AND APPLICATIONS OF THE SAME
: FILE REFERENCE: 266426USOXCIP
: CURRENT APPLICATION NUMBER: US/11/065,943
: CURRENT FILING DATE: 2005-02-25
: PRIOR APPLICATION NUMBER: US 10/787,219
: PRIOR FILING DATE: 2004-02-27
: NUMBER OF SEQ ID NOS: 106
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 24
: LENGTH: 562
: TYPE: PRT
: ORGANISM: Thermus aquaticus
: US-11-065-943-24

```

Query Match 4.7%; Score 76; DB 7; Length 562

Best Local Similarity 41.7%; P-vec. NO. 6.3;
Matches 70; Conservative 39; Mismatches 136; Indels 74; Gaps 14.

```

QY      4  VOALMPBEAIIETLLCYEYL-----HPHMLLELATTYTCRAN----- 42
      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      242 LEALNAHPHYVKIKTIQYRELNLKSKQITQDPLHPR-----IGRLHTRNQATQT 294
      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      43  --CBGPAQLOALNARCPFLAVCLAQQLPEDEPQSSSVFEMVKLVDSMGVELASVRA 100
      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      235 GRLLSQPFLQNI PVR-TPLGQRIIRTFIAESRLLVALVDNYIEL-----RV 341
      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

| | | | |
|--|-----|---|-----|
| QY | 101 | LCOLQWMDHEPRTRGVARGVGLVFESEBIAFHL-RSPGDLTAEBKDOI CDF--LYGRVQARE | 157 |
| Db | 342 | LAHLSDGENDLIRVPEBGKDIHETLSMWTGCVREAVDPLMKRAAKTINGVLTGMSAHL | 401 |
| QY | 158 | ROALARLRTRTOAF--HSAVAPSCGPCLEQODENKSTK-LKDLIG--RYEEBEGQEP | 210 |
| Db | 402 | SOELALIPYEAQAFFERYFQSEFPKVRAMITKLTLEBGRRGVYETLLGRRRYLPLDLAQVK | 461 |
| QY | 211 | GGMDAQC-----PEPGCA-----RLDQMEQVRCDIQFPLSLR-PEEKSS | 251 |
| Db | 462 | NVREAAERARAFPMPOGTRADIMKLAMVTLFPLREMGARMLLQVHDEVLTAAPKE--GA | 519 |
| QY | 252 | RAVARIFPHIGSPCYPAQV | 270 |
| Db | 520 | EAVARLAKEMVEGVYPLAV | 538 |
| RESULT 8 | | | |
| US-10-995-561-943 | | | |
| : Sequence 943, Application US/10995561 | | | |
| : Publication No. US20050272054A1 | | | |
| GENERAL INFORMATION: | | | |
| : APPLICANT: CARGILL, Michele et al. | | | |
| : TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH | | | |
| : TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF | | | |
| : FILE REFERENCE: CLO01559 | | | |
| : CURRENT APPLICATION NUMBER: US/10/995,561 | | | |
| : CURRENT FILING DATE: 2004-11-24 | | | |
| : NUMBER OF SEQ ID NOS: 85702 | | | |
| : SOFTWARE: FastSeq for Windows Version 4.0 | | | |
| : SEQ ID NO 943 | | | |
| : LENGTH: 662 | | | |
| : TYPE: PRT | | | |
| : ORGANISM: Homo sapiens | | | |
| : US-10-995-561-943 | | | |

```

RESULT 9
US-10-995-561-942
: Sequence 942, Application US/10995561
: Publication No. US20050272054A1
:
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
: TITLE OF INVENTION: DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01559
: CURRENT APPLICATION NUMBER: US/10/995,561
: CURRENT FILING DATE: 2004-11-24
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 942
: LENGTH: 702
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-995-561-942

```

```

Query Match      4.7%: Score 75.5; DB 6, Length 702;
Best Local Similarity 27.1%: Pred. No. 9.3;
Matches 36; Conservative 16; Mismatches 58; Indels 23; Gaps 5

QY 21 LEIAPHHMLLIATYTHCRINCPGPAOL---QALAHRCPLVAVCLAQULPEPDG---- 73
Db IEVQVSDWLIRFASIGSFSLITLVGLVGLGNRAARHLCPPLTPPCASSAIEFGGKET 638
QY 74 -QGSSSVSEFDNVKLVDSMGWELAVRRAL-CQLQMDHEPRPGVARGTGVVVEFSGLAH 131
Db WQWNPVDFQ-----EASLQELVWEMSMWDKGERTEPLEKT-----ELPEGAPEL 684
QY 132 RSPGDLTAEEKDQ 144
Db ALDTFELSDQDR 697

```

```

RESULT 10
US-10-995-561-941
: Sequence 941. Application US/10995561
: Publication No. US20050272054A1
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
: TITLE OF INVENTION: DETECTION AND USES THEREOF
: FILE REFERENCE: C1001559
: CURRENT APPLICATION NUMBER: US/10/995,561
: CURRENT FILING DATE: 2004-11-24
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 941
: LENGTH: 754
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)...(754)
: OTHER INFORMATION: Xaa = Any Amino Acid
US-10-995-561-941

```

RESULT 11
US-11-000-463--897
; Sequence 897, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Weinman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping

APPLICANT: Cao, Yi-Cheng
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIP4CN
CURRENT APPLICATION NUMBER: US/11/000,463
CURRENT FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 897
LENGTH: 366
TYPE: PRT
ORGANISM: Homo sapiens
US-11-000-463-897

Query Match 4.6%; Score 75; DB 7; Length 366;
Best Local Similarity 20.3%; Pred. No. 4.7;
Matches 46; Conservative 35; Mismatches 86; Indels 60; Gaps 8;

QY 54 AHRCPLAVCLAQQLPEPDCGSSSVFEDMKLVDSMGHELASVRALCOLQMDHEPRFG 113
DB 82 APRLPQDPVGMRRQQLQEELEEVYKARLPYMAEAEHVLGMNLGRLQQLPYTMDLMEQVA 141
QY 114 VARGGVLVFSEELAFHLRSPGDLTAE-----KDQICDPLVGRVOAREQALRLRTFQ 169
DB 142 LR-----VOLEQLRVGEDTKAQLGGVDAMALLQG-LQSRVYVHNGRFKEFLH 192
QY 170 AF-----HSVAFPS-----C-----GPCLEQODEERSTR 193
DB 193 PYAESLVSGIGHVQELHRSVAPHAPASPARLSRCVQLSRKLTAKAKLHARIQNLQ 252
QY 194 LKDLGRYFEESGQEPGEMDAGPEPQARLQMDQVRCDIQF 240
DB 253 LRRLSRFAFAGT-----GTERGAGPDP-----QMLSEVRORLOAF 288

RESULT 12
US-11-000-463-425

Sequence 425, Application US/11000463
Publication No. US20050266423A1
GENERAL INFORMATION:
APPLICANT: Tang, Y Tom
APPLICANT: Liu, Chenghua
APPLICANT: Aundul, Vinod
APPLICANT: Chen, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIP4CN
CURRENT APPLICATION NUMBER: US/11/000,463
CURRENT FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25

PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 425
LENGTH: 400
TYPE: PRT
ORGANISM: Homo sapiens
US-11-000-463-425

Query Match 4.6%; Score 75; DB 7; Length 400;
Best Local Similarity 20.3%; Pred. No. 5.2;
Matches 46; Conservative 35; Mismatches 86; Indels 60; Gaps 8;

QY 54 AHRCPLAVCLAQQLPEPDCGSSSVFEDMKLVDSMGHELASVRALCOLQMDHEPRFG 113
DB 116 APRLPQDPVGMRRQQLQEELEEVYKARLPYMAEAEHVLGMNLGRLQQLPYTMDLMEQVA 175
QY 114 VARGGVLVFSEELAFHLRSPGDLTAE-----KDQICDPLVGRVOAREQALRLRTFQ 169
DB 176 LR-----VOLEQLRVGEDTKAQLGGVDAMALLQG-LQSRVYVHNGRFKEFLH 226
QY 170 AF-----HSVAFPS-----C-----GPCLEQODEERSTR 193
DB 227 PYAESLVSGIGHVQELHRSVAPHAPASPARLSRCVQLSRKLTAKAKLHARIQNLQ 286
QY 194 LKDLGRYFEESGQEPGEMDAGPEPQARLQMDQVRCDIQF 240
DB 287 LRRLSRFAFAGT-----GTERGAGPDP-----QMLSEVRORLOAF 322

RESULT 13

US-10-131-826A-136
Sequence 136, Application US/10131826A
Publication No. US20050245730A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresin, Naureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvarcoff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P333ORIC128
CURRENT APPLICATION NUMBER: US/10/131,826A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115

PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 136
LENGTH: 278
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-826A-136

Query Match 4.6%; Score 74; DB 6; Length 278;
Best Local Similarity 22.5%; Pred. No. 4.1;
Matches 64; Conservative 34; Mismatches 86; Indels 100; Gaps 15;

QY 9 MPEAEITLICYELHHPHMLLATTYTHCLNCPG-GPAQLOA-----LAHRCPL 60
DB 4 MPEASRL-----LPLLLLL-----LPAELGPSQAGABENDWVLPBKC--- 49
QY 61 AVC--LAQQLP--EDPG-----QSSSVFPMVKLVDSMGWELASVRAL 101
DB 50 EVCKYVAVELKSAFEFKTEKTEVIGTGYGLDQKASGVKTKSDL-----RLIEVETI 103
QY 102 COLQMD--HEPRG-----VRGTGLVER-----SELAFHL 131
DB 104 CKRLDYSILHKERTGSNRFPAKMSFTFELHNLVHKGVKVVMDIPYELMNETSAEVA-DL 162
QY 132 RSPGDLTAEEBDOQCDPLVGRVQARERQALARLRTPOAFHSVAPFSCGFCLEQQ----- 186
DB 163 KKQCDLVVEBEVEIEDWYRNHQBED-----LTFELCANHVKXGKDTSCILAEQMSGKK 215
QY 187 -----DEERSTRKDLIGRYFEEBGOEPGCMEDAQPE 220
DB 216 GDTAALGGKSKKSSRAKAGRSSSSKKQKELGLEGDPSPF 259

RESULT 14
US-10-821-234-1658
Sequence 1658, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmani, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Precipita
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PC_SEQ_genes Version 1.0
SEQ ID NO 1658
LENGTH: 777
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1658

Query Match 4.6%; Score 74; DB 6; Length 777;
Best Local Similarity 23.7%; Pred. No. 15;
Matches 67; Conservative 35; Mismatches 93; Indels 88; Gaps 16;

QY 11 BEAEITLICYELHHPHML-----LE--LATTYTHCLNCPGPAOL-QALAHRC 57
DB 157 EVAISVSTWLASHPEDFGSEAKQDLRLBSFLQTGVAACK-GVGGASALINLRSV 215
QY 58 PPLAVCLAQ--QLBEDPGQSSSVFPMVKLVDSMGWEL-----ASVRALCOLQ-----W 106
DB 216 DPQAPDLKPLALGDDPADPTDV---LVPLADHIAQLTLDDLEPLNLIPSCLGSLW 272
QY 107 DHEPRTG-----VRGTGLVFEFSLAFHLS-----PGDLT----- 138
DB 273 GHRDRPGHSLCPGVR---ATVTQFNKVAAGAVSVLGATGTGPGGVITIRPLRPQRA 329
QY 139 -----AEKQICDF--LYGRVQARERQALARL-----TFQAFHSVAPFSC 179
DB 330 RLEKMWIRVAECCLLNPFSSVAVVASALQSPHRLRAAWGEATRDLSRVFSSL----- 384
QY 180 GPCLQEODEERSTRKDLIGRYFEEBGOEPGCMEDAQPEPG 222
DB 385 --CQIFSEEDVYSGSRELL---VQSVLQSLFEPHSKKAPRSG 422

RESULT 15
US-10-995-561-773
Sequence 773, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: Cargill, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 773
LENGTH: 3803
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-773

Query Match 4.6%; Score 74; DB 6; Length 3803;
Best Local Similarity 22.2%; Pred. No. 1e+02;
Matches 48; Conservative 34; Mismatches 72; Indels 62; Gaps 10;

QY 52 ALAHRCPLAVCLAQULPEDPGQSSSVFPMVKLVDSMGWELASVRALCOLQMDHEPR 111
DB 2032 SLAERSSILQKAIQ-----SQSVQESLBSLQSIG-----EVEQNLQEK 2071
QY 112 TGVRRGTGLVFEFSLAFHLSFGDLTAER-----KDQICDPLVGRVQARERQALARL 165
DB 2072 QVSSLSGVTIOE--ALATNMKLDQDIARQKSLERATEMTYRFMEATDSTTAVALQKLA 2129
QY 166 RTPQAFHSVAPFSCGFCLEQODEERSTRKDL--GRYFEEBGOEPGCMED-----AQG 218
DB 2130 EVSGRFQUL-----CLQQQEKSS--LKXDLRQAMFELHLSKLDQGFENKSRMLASG 2180
QY 219 PEPQARLQWEDQVRCDIROF-----LSLRPEK 248
DB 2181 NQPDQ-----DITHFPQIGELNDEMDO 2204

Search completed: December 27, 2005, 22:12:55
Job time : 4 secs


```

QY 444 PSUDPTVLPVLSLSPGSLAETPAEVPAOLBQLGHQARRPGQERAVWILSGISTLVLP 503
Db 239 SSVEDALIAVTEASDENLTGLNLVY-----GYDSFRDQLQAIKIKITLGSSSTMVLP 292
QY 504 TGAKSLICYQCPALVLSRSRSPCLTVSPPLSLMDQVSGPPLCKAKACHSGMTKORE 563
Db 293 TGAKSLICYQCPAMIL-----PGITLVSPVLSLMDQLKPLPSITKGLSSSQAREAT 348
QY 564 SVLQKIRPAOVHVLMPREALVGAAGLPAAQLPPVAPACIDEAHCLSQ----- 612
Db 349 ETLKELKEGILKIVLFVSPERLIVNEFLSMFRMSLSVSLVAVDEAHVSECYICLSFLCS 408
QY 613 W-----SHNRPQCYLRV-CYKLRKMGVHCPGLTATRTKRSVAVQHLVAEEDP 663
Db 409 WLPALIRSNRNPSPSYRLKASMLPSELKAECLIAMATATMTTLQAVMSLEI-PESTN 467
QY 664 LAGPAPVPTNLHLVSM-----DRDTQALLTLQGRFONLSIIICNRREDTERIALTL 720
Db 468 LIQSQARDNFELSVLSGARRMD--LLIMESPPIKEISIIIVCKFOYETDMISKYL 525
QY 721 RTCLHAAMPVSGSGRAPKTTAEAVHAGMCSERRRRVQRAFMQQLRVVAVATVAFGMGLDR 780
Db 526 RD-----NNINAKGYHSGLPKADRVRIQESFCSNKKIRVVAVATVAFGMGLDX 571
QY 781 PDVAVNLHLGLPSPFESYVAVAGARGQCPAHCHLPFOQGEDELRLRHVADSTDFL 840
Db 572 GDVAIVHIFSVPGMEEVQEIGRARGRLSYCHLPY--DNDYTLKRLSLHSGVDEY 629
QY 841 AVKLVORVPFACCTCTPRPSBOGAVGSRPVKYPQBAQLSHOAPGPRVCGH 900
Db 630 AVGKFLTHVF-----STETKQHE-----KIC-- 650
QY 901 ERALPIQLTVQALDMPBEALITLLCYELAPHHMLLELTATYTHGRLNCPGSPAQLQALA 960
Db 651 --SLVISAQOKPFMKKEVMQITILHLELGRVQYLRMLPOLNICTTAN-----F 697
QY 961 HR---CERPLAVCLAQOLPBDGCGSSSVFEDMVKLVDSMG--WELASVRALCOLQMDH 1014
Db 698 HKTSDAQNTVLYVQSSFNTLAARSATVAALIKSHVQGLHVPDIPVASSICAT-- 754
QY 1015 EPRTVRGTVLVEFSELAFLHRSPPDLTAEEKD-----QICDFLGRVQAR 1062
Db 755 -----IDVLAELQAL-----KEGVYTELDSAFCTYITLSPKKEICLSLHITKML 799
QY 1063 ERQALALRR-TFOAFHSVAFPGSCPLGQDEERSTLKDLLGRYFEEBEGQEPGMD 1121
Db 800 TEIESCKVRKLDIMSSAAVAIVSNTSELSSGAKQTR--SLQSRIFDYFNG-----D 850
QY 1122 AQGEPQOARLQWEDQ---VRCDIRQFLSLRBEKSSRAVARIFHGIGSPCYPAQVY 1177
Db 851 EKCDSPSKA-----TQCAFLRADIKVFLQSNRQAKETPAIARIMGVSPAFPSWV 904
QY 1178 GQDRFRMRKYLHLSFHALVGLATEEL 1203
Db 905 SK-THFMRGRVAVDFRVIMEAAQTEL 929

```

RESULT 2

```

BVECRQ
DNA helicase recQ - Escherichia coli (strain K-12)
N/Alternate names: DNA-dependent ATPase recQ (EC 3.6.1.-)
C/Species: Escherichia coli
C/Date: 31-Dec-1989 #sequence, revision 10-Oct-1997 #text_change 01-Mar-2002
C/Accession: G65186; J05137; A35776; S30712
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: G65186
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA

```

```

A/Residues: 1-610 <BIAT>
A/Cross-references: UNIPARC:UPI00001661F7; GB:AE000458; GB:U00096; NID:92367299; PIDN:AAK
A/Experimental source: strain K-12, substrain MG1655
A/Ritino, N.; Nakayama, K.; Nakayama, H.
Mol. Gen. Genet. 205, 298-304, 1996
A/Title: The recQ gene of Escherichia coli K-12: primary structure and evidence for SOS+
A/Reference number: J05137; MUID:87115164; PMID:3027506
A/Accession: J05137
A/Molecule type: DNA
A/Residues: 1-256, 'A', 258-610 <IR1>
A/Cross-references: UNIPARC:UPI000016F8A
A/Experimental source: strain K12
A/Note: it is uncertain whether residue 1, 3, 8 (all coded by GTG), or 9 (coded by TTG)
R/UmezU, K.; Nakayama, K.; Nakayama, H.
Proc. Natl. Acad. Sci. U.S.A. 87, 5363-5367, 1990
A/Title: Escherichia coli recQ protein is a DNA helicase.
A/Reference number: A35776; MUID:90319113; PMID:2164680
A/Accession: A35776
A/Status: preliminary
A/Molecule type: protein
A/Residues: 4-8 <UMB>
A/Cross-references: UNIPARC:UPI0000172BEC
R/Daniels, D.L.; Plunkett III, G.; Burland, V.; Blattner, F.R.
Science 257, 771-778, 1992
A/Title: Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5 to
A/Reference number: 530660; MUID:92358234; PMID:1379743
A/Accession: S30712
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 'V', 2-256, 'A', 258-610 <DAN>
A/Cross-references: UNIPARC:UPI0000172EBD; EMBL:M87049
A/Genetics:
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992
A/Gene: recQ
A/Map position: 85 min
A/Start codon: GTG
C/Function:
A/Description: involved in the recF recombination pathway; its gene expression is under t
A/Pathway: recF recombination
C/Superfamily: recQ protein; recQ helicase homology
C/Keywords: ATP; DNA binding; hydrolase; nucleotide binding; P-loop; recF recombination
F:49-56/Region: nucleotide-binding motif A (P-loop)
F:144-149/Region: nucleotide-binding motif B
F:146-151/Region: DEAD motif
F:369-402/Domain: recQ helicase homology <RHH>
Query Match 9.8%; Score 631.5; DB 1; Length 610;
Best Local Similarity 37.6%; Pred. No. 2.6e-26;
Matches 164; Conservative 63; Mismatches 134; Indels 75; Gaps 15;
QY 474 EQLGHQAFRPGQERAVWIRILSGISTLVLPPTGAGKSLCYOLPALVSRSPCLTVVSP 533
Db 21 ETEGYQGFRRGQEHRIIDTVLSGRDCLVMPPTGGKSLCYOIPALLNG-----LTVVVSPL 76
QY 534 LSLMDQVSGLP-CLKAACTHSGMTKQRESVLOKTRAAOVHVMVLTMPALVAGAGLPP 592
Db 77 ISLMDQVQDLOANGVAAACINSGTQREQLLEVWTCGTCQIRLLTYAPERLMDNPLH 136
QY 593 AAQLPPVAFACIDEAHCTLSQWSHNPFCYLRVCKVLEERGVCNCFGLTATARTASDV 652
Db 137 LAHNPVLLA-VDEAHCTSQWGHDFRPEYALAQ-LQRPPTLPFPMALTVATADTTQDI 194
QY 653 ACHLVAEEDPLGAPAVPTNLHLSVSMDBDTPQALLTLQGRFONLSI----- 703
Db 195 VRLIG-----LMDPL-----IQIS-SFDRPNRYLM-----EKFKLDQLMRYVQGRG 238
QY 704 ---IYYNRREDTERIALIRTCILHAAMPVSGSGRAPKTTAEAVHAGMCSERRRRVQAF 760
Db 239 KSGIYYCNRSRAKVEDTAAFL-----QSKGISAAAYVHAGLENNVADVQEF 284
QY 761 MGGQLRVAVATVAFGMLGDPDRAVNLHLGLPSPFESYVAVAGARGDQCPAHCHLP- 818
Db 285 QRDLDQIVAVATVAFGKINPNRVFVHPDI PRNIESYQETGAGRGDLPAEAMLFYDP 344

```

QY 819 -----OPQGEDLRELRH---VHADSTDFLAVKRLV-----QRFVPACTC- 855
 Db 345 ADMAMLRCLBEKPPQGO-LQDIERHKLNAMGAFAEAQTCRRLVLTNYGSGRQEPFCGNC 403
 QY 856 TCTRPSEOGAVGGE 871
 Db 404 ICLDPPKQYDGSSTDAQ 419

RESULT 3

ATP-dependent DNA helicase [imported] - Escherichia coli (strain O157:H7, substrain EDL958)
 C.Species: Escherichia coli
 C.Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C.Accession: F66063
 RiPerna, N.Y.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew Miller, L.; Grobbeck, B.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamouzis, K.; Apodaca, Nature 409, 529-533, 2001
 A.Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A.Reference number: AB5480; MUID:21074935; PMID:11206551
 A.Accession: F66069
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-611 <STO>
 A.Cross-references: UNIPROT:Q8X8N1; UNIPROT:Q8FEM6; UNIPARC:UPI000000D08FF; GB:AB005174; A.Experimental source: strain O157:H7, substrain EDL953
 C.Genetics:
 A.Gene: recQ
 C.Superfamily: recQ protein; recQ helicase homology

Query Match 9.8%; Score 631.5; DB 2; Length 611;
 Best Local Similarity 37.6%; Pred. No. 2.8e-26; Mismatches 134; Indels 75; Gaps 15;
 Matches 164; Conservative 63; Mismatches 134; Indels 75; Gaps 15;
 474 BOLGHOAFRPGQERAVMRILSGISTLVLPFGAGKSLCYQLPALLYSRRSPCLTLVVSPL 533
 21 ETFGIGQFRPGQEBITDVLGSDCLVMPFGGKSLCYQIPALLNG---LTVVSPL 76
 QY 534 ISLMDQVSGLP-CLKAACHSGMTRKQRESVLQKIRAAQVHVMLTPEALVGAGLPP 592
 Db 77 ISLMQVQVQLOANGVAACLNSTQTRQLEVMTCGTGQIRLLYIAERLMDNLFLEH 136
 QY 593 AAQLPPVAFACIDEAHCLSGSHNFRPCYLAVCKYLRBMGVHCEGLTATATRTASDV 652
 Db 137 LAHNPVLLA-VDEAHCLSGWHDFRPPYALGQ-LRQRPFTLPFWALTATADDTTRDDI 194
 QY 653 AOHVAEERPDHGPAPVPTNLHLSVMDRTDQALLTLQGRFONLDSI----- 703
 Db 195 VRLIG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKFPDQMLKRYQEOG 238
 QY 704 ---IYCNREDTERIALTLRTCLHAAMVPGSGRAPKTTAEAYHAGKMSRERRVQRAF 760
 Db 239 KSGIYCNRAKVEDTAARL-----QSKGISAAAYHAGLENNVADVQERF 284
 QY 761 MGOQLAVVAVAFMGGLDRPDVAVLHGLPPSFESVYQAVGRAGDQPAHCHFL-- 818
 Db 285 QRDDIQIIVAVAFMGINKPNVRFVHDIIPRNISYYOETGRAGDGLPAEAMLYDP 344
 QY 819 -----OPQGEDLRELRH---VHADSTDFLAVKRLV-----QRFVPACTC- 855
 Db 345 ADMAMLRCLBEKPPQGO-LQDIERHKLNAMGAFAEAQTCRRLVLTNYGSGRQEPFCGNC 403
 QY 856 TCTRPSEOGAVGGE 871
 Db 404 ICLDPPKQYDGSSTDAQ 419

RESULT 4

H91222
 ATP-dependent DNA helicase [imported] - Escherichia coli (strain O157:H7, substrain R1WD)
 C.Species: Escherichia coli
 C.Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C.Accession: H91222
 RiHayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasarara, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. Gena Res. 8, 11-22, 2001
 A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom A.Reference number: A96629; MUID:21156231; PMID:11258796
 A.Accession: H91222
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-611 <HAV>
 A.Cross-references: UNIPROT:Q8X8N1; UNIPROT:Q8FEM6; UNIPARC:UPI000000D08FF; GB:BA000007; A.Experimental source: strain O157:H7, substrain R1WD 050952
 C.Genetics:
 A.Gene: ECE4752
 C.Superfamily: recQ protein; recQ helicase homology

Query Match 9.8%; Score 631.5; DB 2; Length 611;
 Best Local Similarity 37.6%; Pred. No. 2.8e-26; Mismatches 134; Indels 75; Gaps 15;
 Matches 164; Conservative 63; Mismatches 134; Indels 75; Gaps 15;
 474 BOLGHOAFRPGQERAVMRILSGISTLVLPFGAGKSLCYQLPALLYSRRSPCLTLVVSPL 533
 21 ETFGIGQFRPGQEBITDVLGSDCLVMPFGGKSLCYQIPALLNG---LTVVSPL 76
 QY 534 ISLMDQVSGLP-CLKAACHSGMTRKQRESVLQKIRAAQVHVMLTPEALVGAGLPP 592
 Db 77 ISLMQVQVQLOANGVAACLNSTQTRQLEVMTCGTGQIRLLYIAERLMDNLFLEH 136
 QY 593 AAQLPPVAFACIDEAHCLSGSHNFRPCYLAVCKYLRBMGVHCEGLTATATRTASDV 652
 Db 137 LAHNPVLLA-VDEAHCLSGWHDFRPPYALGQ-LRQRPFTLPFWALTATADDTTRDDI 194
 QY 653 AOHVAEERPDHGPAPVPTNLHLSVMDRTDQALLTLQGRFONLDSI----- 703
 Db 195 VRLIG-----LNDPL-----IQIS-SFDRPNIRYMLM-----EKFPDQMLKRYQEOG 238
 QY 704 ---IYCNREDTERIALTLRTCLHAAMVPGSGRAPKTTAEAYHAGKMSRERRVQRAF 760
 Db 239 KSGIYCNRAKVEDTAARL-----QSKGISAAAYHAGLENNVADVQERF 284
 QY 761 MGOQLAVVAVAFMGGLDRPDVAVLHGLPPSFESVYQAVGRAGDQPAHCHFL-- 818
 Db 285 QRDDIQIIVAVAFMGINKPNVRFVHDIIPRNISYYOETGRAGDGLPAEAMLYDP 344
 QY 819 -----OPQGEDLRELRH---VHADSTDFLAVKRLV-----QRFVPACTC- 855
 Db 345 ADMAMLRCLBEKPPQGO-LQDIERHKLNAMGAFAEAQTCRRLVLTNYGSGRQEPFCGNC 403
 QY 856 TCTRPSEOGAVGGE 871
 Db 404 ICLDPPKQYDGSSTDAQ 419

RESULT 5

AH0917
 ATP-dependent DNA helicase (BC 3.6.1.-) [imported] - Salmonella enterica subsp. enterica A/Note: This species has also been called Salmonella typhi
 A/Note: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar A/Reference number: AB0502; MUID:21534947; PMID:11677608
 C.Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C.Accession: AH0917
 RiParthill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, Th. T.; Comerion, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulle, S.; O'Gaora, P. Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A.Accession: AH0917
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-609 <PAR>
 A.Cross-references: UNIPARC:UPI000005A56B; GB:AL513382; PIDN:CAD07934.1; PID:g16504479; C/Genetics:

A:Gene: recQ
C:Superfamily: recQ protein; recQ helicase homology
C:Keywords: hydrolase

Query Match 9.5%; Score 609.5; DB 2; Length 609;
Best Local Similarity 37.7%; Pred. No. 4.1e-25;
Matches 162; Conservative 58; Mismatches 137; Indels 73; Gaps 14;

Qy 474 EQGLHQAFRPQGEAAVMRIISGISTLLVPTGAGSKSLCYOLPALLXSRSCITLVYSP 533
Db 19 ETFGYOQFRPQGEAIIITDALSGRDCLVMPFGGGSLSYQIPALLDQ---LTVVVSPL 74

Qy 534 LSLMDQVDSG-LPPECLKAACHSGMTRKQKQESVLOKIRAAOVHVMILTPREALVGAGLP 592
Db 75 ISLMKDQVDOLANGVAACINSGQSLQEVAGCTGQIRLLIYIAPERLMDNFIDH 134

Qy 593 AAOPLPVAFACTIDEAHCLSQMSHNPFCYLKVCVLRMGVHCFGLTATATRTASDV 652
Db 135 LAHNPVLLA-VDBAHCTISQWGHDFRFEYALGQ-LRQRPALPFPALTATADPTTQDI 192

Qy 653 AQHLVAEERPDHGPAPVPTMLHLSVSMRDTDOALLTLQGRFQNDST----- 703
Db 193 IRLG-----LNDPL-----IQIS-SFDRPNIRYMLM---EKFKPLDQMRVYQEQRG 236

Qy 704 ---IYCNREDTERIALALRTCLHAAMVPGSGRAFKTAAAYHAGMCSERRRVORAF 760
Db 237 KSGIITCNRAKXVEDTTLRLQ-----SRG-----ISAAATHAGLENAIRADVQEK 282

Qy 761 MGOGLRVVATVAFQMGIDRPDVAVLHLGLPSEFSYQAVAGRGDQPAHCHLFLQ 820
Db 283 QRDLQIVVATVAFQMGINKPNVRFVHFPIPRITESYQETGRAGRDGLPAEAMLFYDP 342

Qy 821 -----QGEDRLRRHVHADSTDFLAKRVF-----QVFPACTC-T 856
Db 343 ADMAMLRCLBEKPAQAQODIERHKLAMGAFPAEQTCRRLVLTNYFGEGRQEPGNCDI 402

Qy 857 CTRPPEQEG 866
Db 403 CLDPPKXYDQ 412

RESULT 6
ATP-dependent DNA helicase (EC 3.6.1.-) [imported] - Yersinia pestis (strain CO92)
A:Accession: A10466
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
A:Accession: A10466
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: A10466
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-610 <KUR>
A:Cross-references: UNIPROT:Q8ZAG8; UNIPARC:UPI00000DCADD; GB:AL590842; PIDN:CAC93301.1; C:Genetics:
A:Gene: recQ
C:Superfamily: recQ protein; recQ helicase homology
C:Keywords: hydrolase

Query Match 9.2%; Score 593; DB 2; Length 610;
Best Local Similarity 33.3%; Pred. No. 3.1e-24;
Matches 169; Conservative 71; Mismatches 172; Indels 96; Gaps 18;

Qy 474 EQGLHQAFRPQGEAAVMRIISGISTLLVPTGAGSKSLCYOLPALLXSRSCITLVYSP 533
Db 19 ETFGYOQFRPQGEAIIITDALSGRDCLVMPFGGGSLSYQIPALLDQ---LTVGTLVVSPL 74

Qy 534 LSLMDQVDSG-LPPECLKAACHSGMTRKQKQESVLOKIRAAOVHVMILTPREALVGAGLP 592
Db 75 ISLMKDQVDOLANGVAACINSGQSLQEVAGCTGQIRLLIYIAPERLMDNFIDH 134

Db 75 ISLMKDQVDOLANGVAGACINSGQSLQEVAGCTGQIRLLIYIAPERLMDNFIDH 134
Qy 593 AAOPLPVAFACTIDEAHCLSQMSHNPFCYLKVCVLRMGVHCFGLTATATRTASDV 652
Db 135 LYQWRP-ALLAVDBAHCTISQWGHDFRFEY-RALGQLKQRFEDLVIALTATADATRGDI 192

Qy 653 AQHLVAEERPDHGPAPVPTMLHLSVSMRDTDOALLTLQGRFQNDST----- 703
Db 193 VRLD-----NLDP-----IQIS-SFDRPNIRYMLM---EKFKPLDQMRVYQEQRG 236

Qy 704 ---IYCNREDTERIALALRTCLHAAMVPGSGRAFKTAAAYHAGMCSERRRVORAF 760
Db 237 KSGIITCNRAKXVEDTTLRLQ-----SRG-----LSVAAATHAGLENAIRADVQEK 282

Qy 761 MGOGLRVVATVAFQMGIDRPDVAVLHLGLPSEFSYQAVAGRGDQPAHCHLFLQ 820
Db 283 QRDLQIVVATVAFQMGINKPNVRFVHFPIPRITESYQETGRAGRDGLPAEAMLFYDP 342

Qy 821 QG-----EDRLRRHVHADSTDFLAKRVQVF-----PACTC-T 856
Db 343 ADMAMLRCLBEKPAQAQODIERHKLAMGAFPAEQTCRRLVLTNYFGEGRQEPGNCDI 402

Qy 857 CTRPPEQEGAVGERPVP-KYRQEAEOULSH-----QAAQPRPVCMGHERALPIQLTV 910
Db 403 CLDPPKXYDGLADQKALSCVYRVGQRFGLIYIVLRGANNQIRRMGHDK-----LSV 457

Qy 911 QALDMPBAITLTLCTELHPHMLL 938
Db 458 YGIGR-----EQTHEWVSVL 473

RESULT 7
ATP-dependent DNA helicase RecQ VC0196 [imported] - Vibrio cholerae (strain N16961 serog)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
A:Accession: B82351
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Charadon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB0305; MUID:20406833; PMID:109525301
A:Accession: B82351
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-620 <HEI>
A:Cross-references: UNIPROT:Q9KVFO; UNIPARC:UPI00000C2C1C; GB:AE004110; GB:AE003852; NID: C:Genetics:
A:Gene: VC0196
A:Map position: 1
C:Superfamily: recQ protein; recQ helicase homology

Query Match 9.1%; Score 581.5; DB 2; Length 620;
Best Local Similarity 33.0%; Pred. NO. 1.3e-23;
Matches 160; Conservative 74; Mismatches 171; Indels 80; Gaps 14;

Qy 456 LGPSQALETPAEYFQALBQGLHQAFRPQGEAAVMRIISGISTLLVPTGAGSKSLCYOLP 515
Db 15 MAESSALFAPBERLH--EYFGYQGFVQGVIEALABRDSLVIIMPFGGSKSLCYOIP 72

Qy 516 ALLYSRRSPCLTLVSPUSLMDQVSGLP-CKAACHSGMTRKQKQESVLOKIRAAOV 574
Db 73 ALVLEG-----VTVISPLISLMKDQVDOLKANGVAACINSGTSLAREBLIATYNNMAGOL 128

Qy 575 HVMILTPREALVGAGLPAPPAQPLPVAFACTIDEAHCLSQMSHNPFCYLKVCVLRMGV 634
Db 129 KLVVSPERVLYAEFIRLSHL-PLAMTAVDEAHCTISQWGHDFRFEYASLGQ-LKQRFN 186

Qy 635 HCFGLTATATRTASDVAAQHLVAEERPDHGPAPVPTMLHLSVSMRDTDOALLTLQ 694
Db 187 VPVVALTRITADATRHIMQRLQNEPHQYLSGSDRPNIRYMLVKKKIPVQ-VIRLTET 245

[illegible]

RESULT 8

DNA helicase y1381[imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: H97365
R:Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Quirillo, B., Goldman, A., Liu, F., Mollan, C., Allinger, M., Doughty, D., Scott, C., Leppes, C., Markelz, B. Science 294, 3223-3238, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
R:Reference number: A57359; MUID:21608551; PMID:11743194

A:Accession: A7302
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-602 <KDR>
A:Cross-references: UNIPROT:Q8UJ77; UNIPARC:UPI00000D16A7; GB:AE007869; PIDN:AAK85881.1
C:Genes: AGR_C_92
A:Map position: circular chromosome
C:Superfamily: recQ protein; recQ helicase homology

Query Match 8.7%; Score 560.5; DB 2; Length 602;

Matches 151; Conservative 73; Mismatches 152; Indels 61; Gaps 17;

QY 466 PAEVEQALEQLGHQAFRPGQERAVMRILSGISTLLVLP TGAGKSLCYQLPALYSRRSPC 5255

Db 5 PLNLIKAV--YGYDAFRGRQGEIIQHVVAGNNAFVLMPTGGKSLCYQIPAL--AREG-- 58

QY 526 LTLVSPLLSMDQVSGLPCC-LKACIHSGMTRKQRESVLQKIRAAQVHVLMTPEAL 584

Db 59 MGLVSPILALMVDQVAALRQAGVRAEALNSDLSPEERTLMRDVRAGNVNDILYAAPETL 118

QY 585 VGAGGLPPAAQLPVAFACIDEAHCLSQWSHNFRCYLKRVCKLKRERMGVHCFGLTATA 644

Db 119 LKPDVL-DALQSIDLSLIAVDEAHCLSQMGHDFRPY-RQLDMLIARFPNTPRMALTATA 176

645 TRRTASDVAQH LA VABE PDLHGPA RVPTNLHL SVSMDRDI DQALL I L L QSKRFQN --- LD 701

DD I / DEFIKAEI LGHAI - DEADAF IAGF DRFN I K I A I - M E W N F K I Q L ----- N R I U N G R E N B 223

27

[illegible]

336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

815 ЧИТОВАНИЕ БЕЗ ВЪНТА ДЕТЕТИ. ВЪКЪИ. ОУВЕДАСТСЯ.

Dd 336 DIALRRKFLFE--SDAADQKRYMERQKLDALLGLAETAGRCROYLVSIFEDRCPCGNC D 339
Qy 856 TCTRPSPSEQEAGVGER 872
 ||: || |||: :
Dd 394 TCSSPDLFEGAIQAQK 410

RESULT 9

ATP-dependent DA helicase recQ [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence revision 11-Jan-2002 #ext_change 09-Jul-2004
C/Accession: AB2583
R/wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
C/wood, D.W.; Gillet, W.; Grant, C.; Guenether, D.; Kutayavin, T.; Levy, R.; Li, M.; McClelland
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tso, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, L.
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193

A.Accession: AH2583
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-602 <K1R>
A.Cross-references: UNIPROT:Q8UJ77, UNIPARC:UPI00000D16A7, GB:AE006688, PIDN:AAL41086.1,
A.Experimental source: Strain C58 (Dupont)

C;Genetics:
A;Gene: recQ
A;Map position: circular chromosome
C;Superfamily: recQ protein; recQ helicase homology

Query Match 8.7%; Score 560.5; DB 2; Length 602;

Matches 151; Conservative 73; Mismatches 152; Indels 61; Gaps 17

466 PABVFQAL EQLGHQA FRPGQERAVMRILSGISTLLVLPTGAGKSLCYQLPALLYSRSPC 521

Db 5 PLNLIKAV--YGYDAFRGRQGEIIQHVVAGNNAFVLMPTGGGKSLCYQIPAL--AREG-- 58

526 LTLVSPILSLMBDDQVSGGLPPC-LKACIHSGMTRKQRESVLQKIRAAQVHVLMLTPEAL 58

db 59 MGLVSP LIALMBQVAALRQAGVRAEALNSDLSPEEKRI LWRDVKAGNV DLEIAPEIL II

585 VGAGGLPEAQLPVAFACIDDEAHCLLSQWSHNFKCLIRVCAVDKENIGVHCLDGLAIA 84

DB 119 UNPDAV-DPAHQSIDUSPIAVDEAHLCPSQWGHDFKEFI-KQJDDMUJTHAKFENAEVUNJHJHIAH 17

xy

17. **DEPARTMENT OF THE ARMY**

— — — — —

THE

[illegible]

ИЗДАТЕЛЬСТВО «НАУКА» - МОСКВА 1985

336 DTAENETEE--SDAADOKYMEBOKI.DAII.GIAETAGCBROVI.I.SYEGDRCCEPCGNCD 39

856 TCTRPSPROEGAVGGER 872

Db 394 TCSSPPDLFEGAI AOK 410

RESULT 10

ADJ419
ATP-dependent DNA helicases homolog 1mo2757 [imported] - *Listeria monocytogenes* (strain C:Species: *Listeria monocytogenes*)

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD1419
R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fehli, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A./Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1419
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-590 <GUA>
A/Cross-references: UNIPROT:Q8Y3S4; UNIPARC:UPI00000552D7; GB:NC_003210; PIDN:CAD00970.1
A/Experimental source: strain ESD-e
C/Genetics:
A/Gene: lmo2757
C/Superfamily: recQ protein; recQ helicase homology

Query Match 8.6%; Score 552; DB 2; Length 590;
Best Local Similarity 36.4%; Pred. No. 4.4e-22;
Matches 132; Conservative 66; Mismatches 125; Indels 40; Gaps 9;

QY 474 EOLGHOAFRPGQERAVNRILSGISTLVLPTGAGKSLCYOLPALVSRSPCLTLVSP 533
DB 11 QNFGYQDFRQGVIVISKLSAGEDTLAIMPFGGKSLCYQIPALLFDG---LTIIVSPL 66
QY 534 LSLMDQVSGL-PPCLKAACIHSGMTRKQRESVLQKIRAAQVHVMLTPEALVAGGLPP 592
DB 67 ISLMKQDVDAVLSGIAATPINSITLNRIDIRLDAANSGLKMLYIAPEKILTFPGRL 126
QY 593 AAQIPPAFACIDEAHCLSQSHNFRPCYLAVCKVLRMGVHCFGLTATRTASDV 652
DB 127 IEQVPSLFA-IDEAHCISQWGHDFRPSYTLCDSLDTRRPVIALTATATQAVSDI 185
QY 653 AOHVAABEEDPLHGPAPVPT-----NLHLSVMDRDTQALLTLQGRFQNL-DSITTY 706
DB 186 CRLLKIRAD-----SVKTFPSRDNLAPQVKKGDQKQKYLIDYLT---KNVTSGLTY 235
QY 707 CNREEDTERIALLRICLHAAMVPGSGRAPKTTAE--YHAGMCSRERRRQRAFMQ 764
DB 236 ASTKEVERLRHSPLQ-----KKGVESGMHGMMDLARKMQEFTLYDD 279
QY 765 LRVVATVAFGKGLDRPVRVVLHLGLPPSFESYQAVGARGDQPAHCHLPLQPGED 824
DB 280 IRVIVATVAFGKINKSNVRFVHINIPRINTEAVYQEAARGDGVPSDCILTFSPQSR 339
QY 825 LRE 827
DB 340 IQQ 342

RESULT 11
AE1832
ATP-dependent DNA helicase [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AE1832
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saamoto, S.; Matanabe, A.; Iriyuchi
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AE1832
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-718 <KUR>
A/Cross-references: UNIPROT:Q82093; UNIPARC:UPI000000CDC81; GB:BA000019; PIDN:BA077729.1;
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: alr0205

C/Superfamily: recQ protein; recQ helicase homology

Query Match 8.6%; Score 551; DB 2; Length 718;
Best Local Similarity 35.6%; Pred. No. 6.2e-22;
Matches 153; Conservative 71; Mismatches 148; Indels 58; Gaps 17;

QY 471 QALE-QGHOAFRPGQERAVNRILSGISTLVLPTGAGKSLCYOLPALVSRSPCLTLV 529
DB 9 QALKYHFGYDIFRFGQRIEDALQNRDLVVMPTGGKSLCFQIPALM----KQGLTV 64
QY 530 VSPPLSLMDQVSGL-PPCLKAACIHSGMTRKQRESVLQKIRAAQVHVMLTPEALVAG 588
DB 65 VSPPLALMDQVSLRRNNISATLNSLNAVQVRSREBALINKVLLVYAPRSLSER 124
QY 589 GLP---PAAQIPPAFACIDEAHCLSQSHNFRPCYLAVCKVLRMGVHCFGLTATAT 645
DB 125 FLPLDLVKEKVGISIPALDEAHCVSEMGHDFREY-RQKSLAKRVPVNLATVAT 183
QY 646 RRTSDVAQHVAABEEDPLHGPAPVPTNLHLSVMDRDTQALLTLQGRFQNLDSITTY 705
DB 184 DRVADITIQGL-KQPSIHLSAFNRQNLVYER--PKSQVAVALLELR-DNEGSTII 239
QY 706 YCNREEDTERIALLRICLHAAMVPGSGRAPKTTAEYHAGMCSRERRRQRAFMQ 765
DB 240 YCLTRKVEELTFPL-----QKDKLSALSYHAGLPDDRSKKQTRFIRDV 285
QY 766 RVVATVAFGKGLDRPVRVVLHLGLPPSFESYQAVGARGDQPAHCHLPL----- 818
DB 286 RNVATVAFGKINKPDRVRLVHFDIPRLNESYQESGARGDQPAHCHLPL----- 345
QY 819 -----QPGEDL--RELRRHV-HADSTPLAVKRLVQ-----RVFPA-C-TCT 858
DB 346 IEWSIEKQTPQEDLAKQQLRQMDVYABGT---CRRTIQGVGFRFGNCGNCR 402
QY 859 RPSRQCAV 868
DB 403 YPKMQDWTI 412

RESULT 12
AE1794
ATP-dependent DNA helicases homolog lin2900 [imported] - *Listeria innocua* (strain Clip1);
C/Species: *Listeria innocua*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AE1794
R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fehli, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A./Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AE1794
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-590 <GUA>
A/Cross-references: UNIPROT:Q92620; UNIPARC:UPI000000CC9F3; GB:AL592022; PIDN:CAC98125.1;
A/Experimental source: strain Clip11262
C/Genetics:
A/Gene: lin2900
C/Superfamily: recQ protein; recQ helicase homology

Query Match 8.5%; Score 548; DB 2; Length 590;
Best Local Similarity 36.5%; Pred. No. 7.2e-22;
Matches 132; Conservative 65; Mismatches 127; Indels 38; Gaps 9;

QY 474 EOLGHOAFRPGQERAVNRILSGISTLVLPTGAGKSLCYOLPALVSRSPCLTLVSP 533
DB 11 QNFGYQDFRQGVIVISKLSAGEDTLAIMPFGGKSLCYQIPALLFDG---LTIIVSPL 66
QY 534 LSLMDQVSGL-PPCLKAACIHSGMTRKQRESVLQKIRAAQVHVMLTPEALVAGGLPP 592

Db 67 ISLMKDQDALVSDGIATFTNSTLTNREIDIRLDAFSGELKMLYAPIERIETPGFORL 126
 QY 593 AAOLPVPFACIDBAHCISQWSHNPFCYLRCVKYLBEMGVHCIFGLTATATRTASDV 652
 Db 127 IEQVPISTFPA-IDEAHCISQWGHDFRPSYLSICDSDLDKTRRPVIALITATQAVSDI 185
 QY 653 AQHLVAEERPDHGPAPVPT-----NLHLVSMDRDTQALLTLLOGRFONLSIIYC 707
 Db 186 CRLKI-KPN-----SVIKTGFSRNLAFOVVKGGDKKYLIDYLTG--NSTSGIIYA 236
 QY 708 NREEDTERIAALRTCLHAAMVPGSGRAPKTTAA--YHAGMCRERRRRVRAWQGL 765
 Db 237 STRKEVERHISFTL-----KKGVESGMVHGMDTDIARDWQEKFLYDII 280
 QY 766 RVVAVTVAFCMGDLDPDRAVLHLGLPSPFESYVAVAGRAGRDGQPAHQHLELOFQEDL 825
 Db 281 RYIVATNAFPGMGINRNVAFVIHINIPIRYEATVYQBAGRAGDVPSCDILFLSPQDSRI 340
 QY 826 RE 827
 Db 341 CQ 342

RESULT 13

ATP-dependent DNA helicase RecQ PA3344 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: E83226
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Muzoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lam, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: E83226
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-712 <STO>
 A:Cross-references: UNIPROT:Q9HY01, UNIPARC:UPI00000C59C6; GB:AE004757; GB:AE004091; NID
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: recQ; PA3344
 C:Superfamily: recQ protein; recQ helicase homology

Query Match 8.5%; Score 545; DB 2; Length 712;
 Best Local Similarity 28.5%; Pred. No. 1.3e-21;
 Matches 221; Conservative 102; Mismatches 294; Indels 158; Gaps 32;

QY 471 QALBOL-----GHQAPRPGQERAVMRILSGISTTLVPTGAGKSLCYQLPALYSRRSPCL 526
 Db 4 QALRIKDVFGDAFPGNARITIERVABEGDALVLMPTGGKSLCYQVPLLRBEG---L 59
 QY 527 TLVAVSLTSLMDQV-----GLPCLKAKAICHGTRKQRESVLQKIAAQVHVLMLTP 581
 Db 60 TVVAVSLTSLMDQVATLDELGP-----AVALNSTINPQGRDIARLRGELIKILYLP 115
 QY 582 EALVAGGLPPAAQLPPVAFACIDBAHCISQWSHNPFCYLRCVKYLBEMGVHCIFGLT 641
 Db 116 ERLVQPRMLAFLQRLTFLVGLFA-IDEAHCISQWGHDFRPSYLSICDSDLDKTRRPVIALITATQAVSDI 173
 QY 642 ATATRTASDVQOHVAEERPDHGPAPVPTNLHLSVMDRDTQALLTLLOGRFONLSIIYC 701
 Db 174 ATDMETREEMIQRLHQLQNAEQFLSSFDPRNITRYLVKQPKQQLGLGFLSKRG---D 229
 QY 702 SIIYCNREEDTERIAALRTCLHAAMVPGSGRAPKTTAAEYHAGMCRERRRRVRAWQGL 761
 Db 230 AGIVYCLSRKRVAEVLEFL-----GNQG-----FRALPYHAGISNELRAHMQKFL 275
 QY 762 QGGLRVVAVTVAFCMGDLDPDRAVLHLGLPSPFESYVAVAGRAGRDGQPAHQHLELOFQ 821
 Db 276 NEBGLIMVATIAFAGMGIDKENVAFVAILDLKSLIAYOETGAGRGDGLPADAMWAVGLQ 335

QY 822 GEDIRELRHVHADSTDFLAVKRVQVFPACTCTCTRPSEEGQANG--GERPVP-- 875
 Db 336 --DVILLRQWQMSSECD-ERHRRVERHKLZAMLLCETRCRRQALALVPEEMQPCGH 392
 QY 876 -----KTPPEAQLSHQAP-----GPRVCMGHERALPQLTVQALDMEBAIETLLC 925
 Db 393 CDNCVDGVEVWDVATESARQALSAIRSGQRYGVH-----LVDIILGRETETKRLS-- 443
 QY 926 YLELHPH-----HW-----LELATYTHCRNCPGPAQLOALAHRCPPV-- 966
 Db 444 ---GHQLAVRGICKRGNEDEMRLLFRQLVARGLADVDLDFGG---LRITLACRPLLR 496
 QY 967 ---AVCLAQQLPEDPGQSS-----SVEFDVYKLVDSKWE-LASVRLCOLQWDHERT 1018
 Db 497 GEVRLLELRDLKPPORAKSSSGGASASQLVRSERREMMEARLRLKRLAE---EHSVP 553
 QY 1019 GVRKGTGVLFVFSHLAFLRLSPDLTBKQKQIDFLYGRVQAREQLARLRTPQAF- 1077
 Db 554 YVPPDATTLEW--LRQPRSLSDMAQ-----VSGVGAR-----KLERYQAF 595
 QY 1078 -----HSVAFPSG-GPCLQQDEERSTRKDLGRVPEEEOGEPGG 1118
 Db 596 DVLTDSPAAPAPFODLRHSLASACAGMTPAQLARQNCSEKVVYAMLAEALINQOVSL 655
 QY 1119 MEDAGPPEPGQARLQDWEQVRCIDRQFL-----SLRP-----EKKFSRAVARIFH 1165
 Db 656 EQALDLPPELLIGEIQD-----AFLEBDGELPVAALBERGKRVPGVLH 700

RESULT 14

ATP-dependent DNA helicase homolog - Bacillus subtilis
 N:Alternate names: ATP-dependent DNA helicase recQ
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: A69691
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertorello, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 380, 249-256, 1997
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69691; MUID:98044033; PMID:9384377
 A:Accession: A69691
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-496 <KUN>
 A:Cross-references: UNIPROT:P50729; UNIPARC:UPI000006066D; GB:Z99115; GB:Z99116; GB:AL001
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: recQ
 C:Keywords: ATP; nucleotide binding; P-loop
 F:38-45/Region: nucleotide-binding motif A (P-loop)
 F:132-137/Region: nucleotide-binding motif B
 F:136-139/Region: DEAH motif

Query Match 8.4%; Score 542; DB 2; Length 496;
 Best Local Similarity 36.8%; Pred. No. 1.2e-21;
 Matches 137; Conservative 68; Mismatches 117; Indels 50; Gaps 12;

QY 471 QALBOL-----GHQAPRPGQERAVMRILSGISTTLVPTGAGKSLCYQLPALYSRRSPCLTV 529
 Db 6 QTLVQPFPGTSPKQDQDILIESISGKDTIAMLPTGGKSLCYQLPGYMLDG---WLLI 61

| | | | |
|----|-----|--|-----|
| Qy | 530 | VSPIILSLMDDDVSGIPPC- LKAACISHGSMTRKQRESLQIKIRAAQVHVLMLTPBALVGA | 587 |
| Db | 62 | VSPIILSLIMDDVOOQIKRAGEKRRAAALNSMLNROEKFLEHIIH- RYKFLYISPEALQSP | 119 |
| Qy | 588 | GGLPAPQAQLPVAFACIDEAHCLSQWSHNPRPCYLAVCKULREMGVHCFLGLTATATR | 647 |
| Db | 120 | YVLEKLSKV- DISLFLVIDEAHCISBGMGDFRPDYSKLQ- LRKLGHPVVALATATATRKE | 177 |
| Qy | 648 | TASDV-----AQLTAAVEEPDLHGPAVPVPTVHLHSVEMDRDTPDALITLLQGRFQ | 698 |
| Db | 178 | TLQGVVMNLLEIGHAVRHLSVNR-----NIALRVENADPTAEKTRVLIQ- LVE | 225 |
| Qy | 699 | NLDSI-IYCNRRREDTERIALRLTCLAAWVPSGGRAPTTAA- YHAGMCSRRERR | 755 |
| Db | 226 | NLGPRGIYYCPTRKMAKELAGEIKS-----KTSRADPTYHGLESGRIL | 270 |
| Qy | 756 | VQRAFMGQGLRVVATVAFGMGLRDPVRAVLHGLPSPESYVQAVGRAGRDQOPAHCH | 815 |
| Db | 271 | IQOQFIHQGLVVICCTNAFGMGVDPKDIRYVYHFLPQTAAFMQETIGRAGRDGPVSVI | 330 |
| Qy | 816 | LFLQPGSEDLRE | 827 |
| Db | 331 | LLRAPGDFELQ | 342 |

RESULT 15

ATP-dependent DNA helicase hus2 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S62467; T38578; T43265
R:Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, October 1995
A:Reference number: S62465
A:Accession: S62467
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1328 <BAD>
A:Cross-references: UNIPROT:00981; UNIPARC:UPI000012CE81; EMBL:Z54354; NID:g1033398; P:R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Raftandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z21745
A:Accession: T38578
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1328 <BA2>
A:Cross-references: UNIPARC:UPI000012CE81; EMBL:Z54354; PIDN:CAA91177.1; GSPDB:GNO0066;
A:Experimental source: strain 972h-; cosmid c2G11
R:Stewart, E.; Chapman, C.R.; Al-Khodairy, F.; Carr, A.M.; Enoch, T.
EMBO J. 16, 2682-2692, 1997
A:Title: Rhl1+, a fission yeast gene related to the Bloom's and Werner's syndrome genes,
A:Reference number: Z07997; MUID:97327548; PMID:9184215
A:Accession: T43265
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1328 <STE>
A:Cross-references: UNIPARC:UPI000012CE81; EMBL:Y09426; NID:g1684753; PIDN:CAA70577.1; F
C:Genetics: hus2; SPAC2G11.12
A:Map position: 1L
C:Superfamily: Bloom's syndrome helicase; recQ helicase homology
C:Keywords: ATP; nucleotide binding; P-loop
F:541-548/Region: nucleotide-binding motif A (P-loop)
F:647-652/Region: nucleotide-binding motif B
F:651-654/Region: DEAD motif
F:875-914/Domain: recQ helicase homology <RHH>

Query Match 8.2%; Score 525.5; DB 2; Length 1328;

Matches 143; Conservative 49; Mismatches 154; Indels 39; Gaps 11;

```

0x      443  VPSLDPTVPLYSLSGSGQLAETPAEVFOALEQLGH-QAFRQGERAVMRILSGISTLLV 501
      :| | | | | : | | | | | : | | | | | : | | | | | :

```

```

Db 491 MSLDD---PMLSTPMS-----KEVIGLCKHKHFHLKGRFKQLEAINGTLSGKDVFIL 540
Qy 502 LPRGAGKSLCYOYRALYLSRSRSPCLTVLVSPSLYSIMDQVSGLP----PCILKAACISGSM 557
Db 541 MPTGGSKSLCYOYRALYLSRSRSPCLTVLVSPSLYSIMDQVSLDLHLKRLNIPSLPLSGEPAD 600
Qy 558 TRKQRESVLQKIRAAQVHVLMLTPEALVAGG---LPRAQLPVPAFACIDEANCLISOW 613
Db 601 ERQVSVFPLM-AKVNLVTKLYVTPPEGLASNGAIVRLKSLYERKILARIYIDEAHCVSHW 659
Qy 614 SHHFRFCYLRVCKVTLERMGVHCIFGLTATLTTRTASVAGHNLVAEERDLHGPAPPTN 673
Db 660 GHDFREDYKOL-GLLRDYYQIIPFMAALTATANEIVKDDIINTLEMENCLBEKSSFNRP-N 717
Qy 674 LHLVSMDPBDTQALTLTLTGKRFQNLSTIIYCNBRDETRIALALRTCLH-AAWPGS 732
Db 718 LPEIETPKDOLTYELTYRFLSNGHLH--ESGIIITCLSTRSCQVAAKIRNDYGLKAW---- 771
Qy 733 GGRAPRTTAEAVHAGMCSRRERRRVQRAFMQGLRVVVATVAFGMLDRPDVRAVTLHLGLP 792
Db 772 -----HHAHGLEKVERQRIRINEMQSSYKILVATTAFGGVGDKGQVRFVLIHHSFP 821
Qy 793 PSFESTVQAVGSRAGDGPRAHCPLP 817
Db 822 KSLLEGYQOETGRAGRDGKPAHCIMF 846

```

Search completed: December 27, 2005, 21:54:05
Job time : 48.6 secs

CC helicase gene and its products, and anti-RecoQ4 helicase antibodies are
CC useful in the diagnosis, especially prenatal diagnosis, and treatment of
CC Rothmund-Thomson syndrome. The genomic sequence may especially be used in
CC gene therapy for this condition

XX Sequence 1208 AA;

Query Match 100.0%; Score 6424; DB 3; Length 1208;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MERLDVNERLOAERARORGRRRPSODDYEAARETRALYREYRTLKRTTGAAGSLR 60
DB 1 MERLDVNERLOAERARORGRRRPSODDYEAARETRALYREYRTLKRTTGAAGSLR 60
QY 61 SSELPAAEAEAPERCWCPHINRAATYSPOPTGRSRQSGVPDYQRLKANLKTTOAG 120
DB 61 SSELPAAEAEAPERCWCPHINRAATYSPOPTGRSRQSGVPDYQRLKANLKTTOAG 120
QY 121 PALGRRPWPLGRASSKASTPKPGTGPVPSFAKXVSDSPOLPEPQPRGRLOHLQASLS 180
DB 121 PALGRRPWPLGRASSKASTPKPGTGPVPSFAKXVSDSPOLPEPQPRGRLOHLQASLS 180
QY 181 QRLGSLDPGWLQCHSEVDFLGAAPKACRPDLGSEESQLLPGESAVLGPAGSGQPEAS 240
DB 181 QRLGSLDPGWLQCHSEVDFLGAAPKACRPDLGSEESQLLPGESAVLGPAGSGQPEAS 240
QY 241 AFQSVSTRVSGPSSSGGGRKRWNEBWPESPAVOQSSSQAGPSSBAGAVAEEDPG 300
DB 241 AFQSVSTRVSGPSSSGGGRKRWNEBWPESPAVOQSSSQAGPSSBAGAVAEEDPG 300
QY 301 EPVQAGPPQPCSSSPNPRYHGLSPSSQARAGKAEGLTAFLHFPRLARHDRGNVYRLNMKQ 360
DB 301 EPVQAGPPQPCSSSPNPRYHGLSPSSQARAGKAEGLTAFLHFPRLARHDRGNVYRLNMKQ 360
QY 361 KHYVGRALRSRLLRKQAMQKMRKKECFGGGATVTTKESCFLEQFDHMAQCPERP 420
DB 361 KHYVGRALRSRLLRKQAMQKMRKKECFGGGATVTTKESCFLEQFDHMAQCPERP 420
QY 421 SEEDTDVAGPPPLVPSPPVPEVPSLDPTVPLYSLSGSGQALTPAFAVFOALRQLGHO 480
DB 421 SEEDTDVAGPPPLVPSPPVPEVPSLDPTVPLYSLSGSGQALTPAFAVFOALRQLGHO 480
QY 481 FRPGQERAVMRLISGISITLVLPFGAGKSLCYOLPALIYSRSPCLTVSPILSIMDDQ 540
DB 481 FRPGQERAVMRLISGISITLVLPFGAGKSLCYOLPALIYSRSPCLTVSPILSIMDDQ 540
QY 541 VSGLPPLCKAACIHSQMTKQRESVLQKIRAAOVHVLMTPEALVGAAGLPPAQLPEVA 600
DB 541 VSGLPPLCKAACIHSQMTKQRESVLQKIRAAOVHVLMTPEALVGAAGLPPAQLPEVA 600
QY 601 PACIDEAHCLSQSHNPRCYLRCKVLRBMGVCFLGTATATRTRTASVAGHLAAVE 660
DB 601 PACIDEAHCLSQSHNPRCYLRCKVLRBMGVCFLGTATATRTRTASVAGHLAAVE 660
QY 661 EPDLHGPAVPVNTNHLISVMDRTDQALTLTLOGKRPONLSIIICMRREBDETRIAL 720
DB 661 EPDLHGPAVPVNTNHLISVMDRTDQALTLTLOGKRPONLSIIICMRREBDETRIAL 720
QY 721 RTCLHAAMVPSGGARPKTTAEAYHAGCSERRRRVQRAFNGQOLRVVATVAFGMLDR 780
DB 721 RTCLHAAMVPSGGARPKTTAEAYHAGCSERRRRVQRAFNGQOLRVVATVAFGMLDR 780
QY 781 PDVAVAHLGLPSPFESTVQAVGARDGQPAHCHLFTQPGEDLRELRHHVHVDSTDFL 840
DB 781 PDVAVAHLGLPSPFESTVQAVGARDGQPAHCHLFTQPGEDLRELRHHVHVDSTDFL 840
QY 841 AVKLVORVPACCTCTRPSEBEGANGERRPVKYPPOABOLSHQAAAGPRVCMGH 900
DB 841 AVKLVORVPACCTCTRPSEBEGANGERRPVKYPPOABOLSHQAAAGPRVCMGH 900
QY 901 ERALPIQLTVQALDMPBEALETLLCYLELPHHMLLELATYTYTCRLNCPGPAQLQALA 960
DB 901 ERALPIQLTVQALDMPBEALETLLCYLELPHHMLLELATYTYTCRLNCPGPAQLQALA 960
```

```
DB 901 ERALPIQLTVQALDMPBEALETLLCYLELPHHMLLELATYTYTCRLNCPGPAQLQALA 960
QY 961 HRCPPPLAVCLAQQLPEDPGQSSSVFEDMYKLVDNMGMEILASVRRALCOLQMDHEPRTGV 1020
DB 961 HRCPPPLAVCLAQQLPEDPGQSSSVFEDMYKLVDNMGMEILASVRRALCOLQMDHEPRTGV 1020
QY 1021 RRGTVLVPSFELAFLRSRPGDLFAERKDOI CDPLVGRVQAREQALRLRRTQAFHSV 1080
DB 1021 RRGTVLVPSFELAFLRSRPGDLFAERKDOI CDPLVGRVQAREQALRLRRTQAFHSV 1080
QY 1081 AFPSGCPLEQODEERSTRKDLGRFYEEBEGQEPGAMEDAOGPEPQOARLODMEQVR 1140
DB 1081 AFPSGCPLEQODEERSTRKDLGRFYEEBEGQEPGAMEDAOGPEPQOARLODMEQVR 1140
QY 1141 CDIRQFLSLRPEEKFSRAVARIFHGIGSPCYAQTGODRRFMRKYLHLSFHALVGLAT 1200
DB 1141 CDIRQFLSLRPEEKFSRAVARIFHGIGSPCYAQTGODRRFMRKYLHLSFHALVGLAT 1200
QY 1201 BELLOVAR 1208
DB 1201 BELLOVAR 1208
```

RESULT 2
ABG93387
ID ABG93387 standard; protein; 1208 AA.

XX ABG93387;
XX 20-NOV-2002 (first entry)

XX Human RecoQ protein-like 4 (RECOQ4).

XX Human; RecoQ protein-like 4; RECOQ4; infection; inflammation; cytostatic;
XX tumour formation; cancer; antiinflammatory; antimicrobial;
XX antisense therapy.

XX Homo sapiens.

XX US6436706-B1.

XX 20-AUG-2002.

XX 23-FEB-2001; 2001US-00792594.

XX 23-FEB-2001; 2001US-00792594.

XX (ISIS-) ISIS PHARM INC.

XX Ward DT, Watt AT;

XX WPI; 2002-689941/74.

XX N-PSDB; ABS68851.

PT New antisense compounds targeted to nucleic acid encoding RecoQ protein-
PT like 4, useful for modulating expression of the nucleic acid and treating
PT diseases associated with expression of the nucleic acid in humans.

PS Disclosure; Col 47-56; 45pp; English.

CC The invention relates to a compound targeted to specific nucleobases of
CC RecoQ protein-like 4 (RECOQ4) and which hybridises and inhibits the
CC expression of RECOQ4. The compound is useful for inhibiting the
CC expression of RECOQ4 in cells or tissues and for treating an animal,
CC particularly a human suspected of having or being prone to a disease or
CC condition associated with expression of RECOQ4. The compound is useful
CC for diagnostics, therapeutics and as a research reagent, e.g.
CC prophylactically to prevent or delay infection, inflammation or tumour
CC formation. This sequence represents human RECOQ4 polypeptide

XX Sequence 1208 AA;
Query Match 100.0%; Score 6424; DB 5; Length 1208;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MERLDVREERLOAMERAFRRORGRPSQDDVAAAEETALYREYTTLKTTGGAGGGR 60
DB 1 MERLDVREERLOAMERAFRRORGRPSQDDVAAAEETALYREYTTLKTTGGAGGGR 60
QY 61 SEESLPAABEABERPCWEPHINRAATKSPPTPGSRGQSVDPGORKANLKGTLQAG 120
DB 61 SEESLPAABEABERPCWEPHINRAATKSPPTPGSRGQSVDPGORKANLKGTLQAG 120
QY 121 PALGRRPWPGLGRASSKASTPKPPGTGVPVFAKVDSEPPOLPEPPRGRLOHLQASIS 180
DB 121 PALGRRPWPGLGRASSKASTPKPPGTGVPVFAKVDSEPPOLPEPPRGRLOHLQASIS 180
QY 181 QRLGSLDPGWLORCHSEVPDPLGAPKACRPDLGSEBSQLLIGESAVLGPAGSGQPEAS 240
DB 181 QRLGSLDPGWLORCHSEVPDPLGAPKACRPDLGSEBSQLLIGESAVLGPAGSGQPEAS 240
QY 241 AROEVSIRVSGPOSSSGGKKRMNEEPSPAOVQOESSQAGPPEEGAGAVAEDEPPG 300
DB 241 AROEVSIRVSGPOSSSGGKKRMNEEPSPAOVQOESSQAGPPEEGAGAVAEDEPPG 300
QY 301 BEVOAQPPOPCSSPNRYHGLSPSSQARAGAKGTAPLHITPRLARHDRGNVYRLNMKQ 360
DB 301 BEVOAQPPOPCSSPNRYHGLSPSSQARAGAKGTAPLHITPRLARHDRGNVYRLNMKQ 360
QY 361 KHYVRGALRSRLRKQAMKQKRRKKGECFGGGAATVTTKESCFLNEOPDMAAOCPRPA 420
DB 361 KHYVRGALRSRLRKQAMKQKRRKKGECFGGGAATVTTKESCFLNEOPDMAAOCPRPA 420
QY 421 SEEDTDVAVPEPLVPSPPQVPEVPSLDPTVPLYSIGPSGQLAETAEVQALBEQGHQA 480
DB 421 SEEDTDVAVPEPLVPSPPQVPEVPSLDPTVPLYSIGPSGQLAETAEVQALBEQGHQA 480
QY 481 FRPGQERAVWRILISGISTLLVLTPTGAGKSLCYQLPALYSRSRSPCLTLVSPSLSLMDQ 540
DB 481 FRPGQERAVWRILISGISTLLVLTPTGAGKSLCYQLPALYSRSRSPCLTLVSPSLSLMDQ 540
QY 541 VSGLPCLRAACIHSQMTKQRESVLQKIRAAQVHVLMTPEALVAGAGLPRAQLPPVA 600
DB 541 VSGLPCLRAACIHSQMTKQRESVLQKIRAAQVHVLMTPEALVAGAGLPRAQLPPVA 600
QY 601 FACIDBAHCLISQNSHNPFCYLAVCVLAERMGVHFLGTATARTSDVAQHLAVAE 660
DB 601 FACIDBAHCLISQNSHNPFCYLAVCVLAERMGVHFLGTATARTSDVAQHLAVAE 660
QY 661 EBDLHGPAPVPTLHLSVSMRDTDOALLTLQGGKFFONLDSIIICNRREDTERIALLL 720
DB 661 EBDLHGPAPVPTLHLSVSMRDTDOALLTLQGGKFFONLDSIIICNRREDTERIALLL 720
QY 721 RTCLHAAMVSGSGRAPKTTAAVYHAGMCSRRRRVQRAFMQGLNVVATVAFMGGLDR 780
DB 721 RTCLHAAMVSGSGRAPKTTAAVYHAGMCSRRRRVQRAFMQGLNVVATVAFMGGLDR 780
QY 781 PDVRAVLAHLGLPSFESSYQAVARAGDQPAHCHFLQGGEDLRLRHVHADSDFL 840
DB 781 PDVRAVLAHLGLPSFESSYQAVARAGDQPAHCHFLQGGEDLRLRHVHADSDFL 840
QY 841 AVKRLVORVPACTCTRPSPSEQGVAGGERVVPKPPPEABQLSHQAPGRRCVMGH 900
DB 841 AVKRLVORVPACTCTRPSPSEQGVAGGERVVPKPPPEABQLSHQAPGRRCVMGH 900
QY 901 ERALPLOTVALDMPBEAETLLCYLELHPHMLBELATTTTHCRINCPCGPAQIQALA 960
DB 901 ERALPLOTVALDMPBEAETLLCYLELHPHMLBELATTTTHCRINCPCGPAQIQALA 960
QY 961 HRCPPLAVCLAQOLPEDPGSSSVEFDWYKLVDSMGWELASVRBALCOLQMDHBPRTGV 1020
DB 961 HRCPPLAVCLAQOLPEDPGSSSVEFDWYKLVDSMGWELASVRBALCOLQMDHBPRTGV 1020
QY 1021 RRGTVLVEFSELAFLHLSPGDLTAEEKQICDPLYGRVQAREQALRLRRTFOAFHSV 1080
DB 1021 RRGTVLVEFSELAFLHLSPGDLTAEEKQICDPLYGRVQAREQALRLRRTFOAFHSV 1080

```

Db 1021 RRGTVLVEFSELAFLHLSPGDLTAEEKQICDPLYGRVQAREQALRLRRTFOAFHSV 1080
QY 1081 AFGSGPCLTEQODERSTRLLDLGRYFEEEGGEPGGMEDAOGEPOQALQMDQDVR 1140
Db 1081 AFGSGPCLTEQODERSTRLLDLGRYFEEEGGEPGGMEDAOGEPOQALQMDQDVR 1140
QY 1141 CDIRQFLSLRPEEKSSRAVARIFHGIGSPCYPAQVYQODRRFWRKYHLHSFHALVGLAT 1200
Db 1141 CDIRQFLSLRPEEKSSRAVARIFHGIGSPCYPAQVYQODRRFWRKYHLHSFHALVGLAT 1200
QY 1201 EBLQVAR 1208
Db 1201 EBLQVAR 1208

RESULT 3
ADQ21562
ID ADQ21562 standard; protein; 1208 AA.
XX
AC ADQ21562;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4382.
XX
KW soft tissue sarcoma; cytoskeletal; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PE 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnick A;
XX
DR WPI; 2004-441208/41.
XX

PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4382; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytoskeletal applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 1208 AA;

Query Match 100.0%; Score 6424; DB 8; Length 1208;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MERLDVREERLOAMERAFRRORGRPSQDDVAAAEETALYREYTTLKTTGGAGGGR 60
DB 1 MERLDVREERLOAMERAFRRORGRPSQDDVAAAEETALYREYTTLKTTGGAGGGR 60

```

```

QY 61 SSESIPAAAEAPBRCWCPHLNRAATKSPOTPRSGSGVDPYGGRLKANTKGLTQAG 120
DB 61 SSESIPAAAEAPBRCWCPHLNRAATKSPOTPRSGSGVDPYGGRLKANTKGLTQAG 120
QY 121 PALGRPFWPLGRASSAKASTPKPPTGTPVPSFAEKYSDPPOLPEBPQPRGLQHLQASLS 180
DB 121 PALGRPFWPLGRASSAKASTPKPPTGTPVPSFAEKYSDPPOLPEBPQPRGLQHLQASLS 180
QY 181 QRLGSLDGMQORCHSEVPDFLGAPKACRPDLGSEBSQLLIPGESAVLGPAGSGGPEAS 240
DB 181 QRLGSLDGMQORCHSEVPDFLGAPKACRPDLGSEBSQLLIPGESAVLGPAGSGGPEAS 240
QY 241 AFOEYSIVGSPOPSSSGGKKRMNMBEPMESPAQVOQSSSOAGPPSEKAGAVAVEDDEPG 300
DB 241 AFOEYSIVGSPOPSSSGGKKRMNMBEPMESPAQVOQSSSOAGPPSEKAGAVAVEDDEPG 300
QY 301 BPVOAQPPQPCSSPSPNRYHGLSPSSOARAGKAGTAPLHFPRLARHRRGNYVLANMKQ 360
DB 301 BPVOAQPPQPCSSPSPNRYHGLSPSSOARAGKAGTAPLHFPRLARHRRGNYVLANMKQ 360
QY 361 KHYVGRALRRLRKQAMQKMRKKGCFGGGGATVTTKSCFLNEQPDHMAQCPRPA 420
DB 361 KHYVGRALRRLRKQAMQKMRKKGCFGGGGATVTTKSCFLNEQPDHMAQCPRPA 420
QY 421 SEEDTDAGPBPVPSPPQVPEVPSLDPTVLPLVSLGPSGQAEPAVFOALBQLGHQA 480
DB 421 SEEDTDAGPBPVPSPPQVPEVPSLDPTVLPLVSLGPSGQAEPAVFOALBQLGHQA 480
QY 481 FRPQGERAVMBILSGISTLLVLPFGAGKSLCYOLPALIYSRRSPCLTIIVSPLSLIMDQ 540
DB 481 FRPQGERAVMBILSGISTLLVLPFGAGKSLCYOLPALIYSRRSPCLTIIVSPLSLIMDQ 540
QY 541 VSGPPLCKAACIHSGMTKQRESVLQKIRAAQVAVMLTPBALVAGAGLPPAAQLPVA 600
DB 541 VSGPPLCKAACIHSGMTKQRESVLQKIRAAQVAVMLTPBALVAGAGLPPAAQLPVA 600
QY 601 FACIDBAHCLSSOMSNFRPCYLKRVCKVLRBMGVCFGLTATARRASDVQAQILAAAE 660
DB 601 FACIDBAHCLSSOMSNFRPCYLKRVCKVLRBMGVCFGLTATARRASDVQAQILAAAE 660
QY 661 EPDLHGPAVPTNLHLVSMDRDTQALLTLQGRFONLDSIIYYCNRREDTERIALLL 720
DB 661 EPDLHGPAVPTNLHLVSMDRDTQALLTLQGRFONLDSIIYYCNRREDTERIALLL 720
QY 721 RTCHHAANVPSSGGAPKTTAAHYAAGCNSRRRRVQDAFMQOQLRVVAVYAFMGILDR 780
DB 721 RTCHHAANVPSSGGAPKTTAAHYAAGCNSRRRRVQDAFMQOQLRVVAVYAFMGILDR 780
QY 781 PDVBAVLHLGLPPSPESYVOAVGRAGRDGPAHCHLFIQOGEEDLRELRHHVHADSTDFL 840
DB 781 PDVBAVLHLGLPPSPESYVOAVGRAGRDGPAHCHLFIQOGEEDLRELRHHVHADSTDFL 840
QY 841 AVKELVORVPACTCTCTRRPSEBOGAVGERPVKYPPOEAEOULSHOAAFGPRRVCMH 900
DB 841 AVKELVORVPACTCTCTRRPSEBOGAVGERPVKYPPOEAEOULSHOAAFGPRRVCMH 900
QY 901 ERAPPIQULTVQALMPEBAITLTLCTYLHHPHMLIELATTYTHCRINCPGGPAQLQALA 960
DB 901 ERAPPIQULTVQALMPEBAITLTLCTYLHHPHMLIELATTYTHCRINCPGGPAQLQALA 960
QY 961 HRCPEPLAVCLAQOQLPEDBGGSSSYEPMMVKLVMSMGELASVVRALQOLQMDHEPRRGV 1020
DB 961 HRCPEPLAVCLAQOQLPEDBGGSSSYEPMMVKLVMSMGELASVVRALQOLQMDHEPRRGV 1020
QY 1021 RRGTVLVEFSELAFLHRSPODLTAEBKDOICDFLYGVARERQALRLRRTFOAFHSV 1080
DB 1021 RRGTVLVEFSELAFLHRSPODLTAEBKDOICDFLYGVARERQALRLRRTFOAFHSV 1080
QY 1081 AFPSGCGFLBQOODEBSRTRLKDLGRYFEEBEGQEPGAMEDAOGPEPQARLQDMEQVVR 1140
DB 1081 AFPSGCGFLBQOODEBSRTRLKDLGRYFEEBEGQEPGAMEDAOGPEPQARLQDMEQVVR 1140

```

```

QY 1141 CDIROPLSRPEEKSSRAVARIFHIGSPCYPAOVYGGDRRFRKYLHLSFHALNGIAT 1200
DB 1141 CDIROPLSRPEEKSSRAVARIFHIGSPCYPAOVYGGDRRFRKYLHLSFHALNGIAT 1200
QY 1201 BELLQVAR 1208
DB 1201 BELLQVAR 1208

RESULT 4
ADU80701
ID ADU80701 standard; protein; 1208 AA.
XX
AC ADU80701;
XX
DT 10-FEB-2005 (first entry)
XX
DE Human RTS-associated protein SEQ ID NO 73.
XX
KW anticancer; apoptosis; RecQ; DNA helicase; RNA interference; RNAi;
KM cytosolic; gene silencing; enzyme.
XX
OS Homo sapiens.
XX
PN WO2004100990-A1.
XX
PD 25-NOV-2004.
XX
PF 19-MAY-2004; 2004WO-JP007145.
XX
PR 19-MAY-2003; 2003JP-00140685.
XX
PA (GENE-) GENE2CARE RES INST CO LTD.
XX
PI Takagi M, Shimamoto A, Furuchi Y, Sato A;
XX
DR WPI; 2004-833899/82.
XX
PT Apoptosis inducing agent of cancer cell useful as anticancer agent.
PT Apoptosis compound that suppresses expression of RecQ DNA helicase-family
PT gene/function of protein encoded by RecQ DNA helicase-family gene, as
PT active ingredient.
XX
PS Disclosure; SEQ ID NO 73; 262bp; Japanese.
XX
CC This invention describes a novel anticancer apoptosis inducing agent
CC which comprises (a) compound, which suppresses the expression of RecQ DNA
CC helicase-family gene, (b) DNA which expresses the double-stranded RNA
CC with RNA interfering (RNAi) effect with respect to RecQ DNA helicase-
CC family gene, or (c) a compound, which suppresses the function of protein
CC encoded by RecQ DNA helicase-family gene, as an active ingredient. The
CC invention also describes a method for screening for a candidate compound
CC with a protein encoded by RecQ DNA helicase-family gene. A compound,
CC which suppresses the expression of RecQ DNA helicase-family gene is a
CC transcription product or antisense nucleic acid of one part of RecQ DNA
CC helicase-family gene, or a nucleic acid which has ribozyme activity which
CC cleaves the transcription product of RecQ DNA helicase-family gene
CC specifically. It can also be a RecQ DNA helicase-family protein variant
CC which has dominant negative character with respect to the protein encoded
CC by RecQ DNA helicase-family gene, an antibody or a low molecular compound
CC coupled with the protein encoded by RecQ DNA helicase-family gene. The
CC RecQ DNA helicase-family gene is a WRN gene, BLM gene or RecQ1 gene. The
CC method of the invention is useful for manufacturing an anticancer agent
CC as a pharmaceutical composition. The novel agent is highly safe,
CC effective and does not cause any side effects when used as a
CC pharmaceutical. The products of the invention have cytostatic activity.
CC This sequence represents a human helicase-associated RTS protein.
XX
SQ Sequence 1208 AA;

```

Query Match 100.0%; Score 6424; DB 8; Length 1208;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MERLDVVERLOAMERAFRRORGRSPSODDVEAABEETRALYREYRTLKRTTGOAGGLR 60
DB 1 MERLDVVERLOAMERAFRRORGRSPSODDVEAABEETRALYREYRTLKRTTGOAGGLR 60
QY 61 SSESIPAAAEABEPRCWPGLNRAATKSPQTPGRSRQSVDPYGOGLKANLKGTLQAG 120
DB 61 SSESIPAAAEABEPRCWPGLNRAATKSPQTPGRSRQSVDPYGOGLKANLKGTLQAG 120
QY 121 PALGRPMPPLGASAKASTPKPGTGPVPSFAEKVSDEBPOLPEOPRGRLOHLQASLS 180
DB 121 PALGRPMPPLGASAKASTPKPGTGPVPSFAEKVSDEBPOLPEOPRGRLOHLQASLS 180
QY 181 ORLGSLDGMLORCSEVPDFLGAPKACRPDLGSEESQLLIIGESAVYAGPAGSQGPBAS 240
DB 181 ORLGSLDGMLORCSEVPDFLGAPKACRPDLGSEESQLLIIGESAVYAGPAGSQGPBAS 240
QY 241 AFOEVSIVGSPQSSSGGKKRMBEPWESPAVOYQOESSQAGPPSEGAVAVEDEPPG 300
DB 241 AFOEVSIVGSPQSSSGGKKRMBEPWESPAVOYQOESSQAGPPSEGAVAVEDEPPG 300
QY 301 BPVQAQPPQPCSSBENPRYHGISSESSQARAGAEGLAHLTPRLARHDKRGYVRLNMQ 360
DB 301 BPVQAQPPQPCSSBENPRYHGISSESSQARAGAEGLAHLTPRLARHDKRGYVRLNMQ 360
QY 361 KHYVGRALRSRLRKQAMKOKMKRKKGCGGGGATVTTKESCFINBOPDHMAACPPRA 420
DB 361 KHYVGRALRSRLRKQAMKOKMKRKKGCGGGGATVTTKESCFINBOPDHMAACPPRA 420
QY 421 SEEDTDVAGPEPLVSPQVPEVPSLDPTVPLVSLGSGQLAETPAVEFQALBQLGHA 480
DB 421 SEEDTDVAGPEPLVSPQVPEVPSLDPTVPLVSLGSGQLAETPAVEFQALBQLGHA 480
QY 481 PRPGQERAVRMLTSGISTLLVPLTGAGKSLCYOLPALYSRSPCLTLVSPSLIMDQ 540
DB 481 PRPGQERAVRMLTSGISTLLVPLTGAGKSLCYOLPALYSRSPCLTLVSPSLIMDQ 540
QY 541 VSGLPPLCAACIHSGMTKORESVLOKIRAAOVHMLTPRALVAGAGLPPAOLPPVA 600
DB 541 VSGLPPLCAACIHSGMTKORESVLOKIRAAOVHMLTPRALVAGAGLPPAOLPPVA 600
QY 601 FACIDBAHCLSQMSHNFPCYLKVCVLERMGVHCFGLTATARTASDVAAOHLAVAE 660
DB 601 FACIDBAHCLSQMSHNFPCYLKVCVLERMGVHCFGLTATARTASDVAAOHLAVAE 660
QY 661 BPDHGPAPVPTNLHLSVSMRDQALITLLIQKKFQMLDSIIICNREDETERIALLL 720
DB 661 BPDHGPAPVPTNLHLSVSMRDQALITLLIQKKFQMLDSIIICNREDETERIALLL 720
QY 721 RCTLAAMWPGSGGAPKTTAAVYHAGMSRRRRYQBARFMOQLAVVVAATVAFGMLDR 780
DB 721 RCTLAAMWPGSGGAPKTTAAVYHAGMSRRRRYQBARFMOQLAVVVAATVAFGMLDR 780
QY 781 PDVRAVTHLGLPSPFESYVQAVGRAGRDQPAHCHFLPOQSEDLRELRRHYADSTDL 840
DB 781 PDVRAVTHLGLPSPFESYVQAVGRAGRDQPAHCHFLPOQSEDLRELRRHYADSTDL 840
QY 841 AYKRLVORVPACTCTCTRPSPSBOGAVGERRVPKYPQOEAOLSHOAPGRRVCMGH 900
DB 841 AYKRLVORVPACTCTCTRPSPSBOGAVGERRVPKYPQOEAOLSHOAPGRRVCMGH 900
QY 901 EERALPTLOLVOLADMEEALETLLCYLEHPPHMLLATTYTHGCLNCGGPAOLQALA 960
DB 901 EERALPTLOLVOLADMEEALETLLCYLEHPPHMLLATTYTHGCLNCGGPAOLQALA 960
QY 961 HRCPLAVCLAQOLPEDPGQSSSVFEDWKLVDNSGWEIASVRRALCOLQMDHEBRTGV 1020
DB 961 HRCPLAVCLAQOLPEDPGQSSSVFEDWKLVDNSGWEIASVRRALCOLQMDHEBRTGV 1020
QY 1021 RRGCTGLVPSFLAFTLBSFGDLTAEKQOICPFLYGRVQARRQALALRRFQAFHSV 1080
DB 1021 RRGCTGLVPSFLAFTLBSFGDLTAEKQOICPFLYGRVQARRQALALRRFQAFHSV 1080

```

```

QY 1081 AFPSGPGCLEQODEENSTRLLKDLIGRYFEEBEGOEPCGMEDAQCPREGQARLODMEDQVR 1140
DB 1081 AFPSGPGCLEQODEENSTRLLKDLIGRYFEEBEGOEPCGMEDAQCPREGQARLODMEDQVR 1140
QY 1141 CDIRQFLSLREBEKFSRRVARIETHIGSPCYPAOVYGGODRFRKYLHLSTFHALVGLAT 1200
DB 1141 CDIRQFLSLREBEKFSRRVARIETHIGSPCYPAOVYGGODRFRKYLHLSTFHALVGLAT 1200
QY 1201 BELLOVAR 1208
DB 1201 BELLOVAR 1208

RESULT 5
ID AAM95050
AAM95050 standard; protein; 1208 AA.
XX
AC AAM95050;
XX
DT 14-MAY-1999 (first entry)
XX
DE Human helicase protein.
XX
KW RecQ4 gene; helicase; Werner's syndrome; Bloom's syndrome; human.
XX
OS Homo sapiens.
XX
PN WO9905284-A1.
XX
PD 04-FEB-1999.
XX
PF 10-JUL-1998; 98WO-JP003114.
XX
PR 25-JUL-1997; 97JP-00200387.
XX
PA (AGEN-) AGENE RES INST CO LTD.
XX
PI Shimamoto A, Kitao S, Furuchi Y;
XX
DR WPI; 1999-142939/12.
XX
DR N-PSDB; AAX21656.
XX
PT New human helicase gene RecQ4 - used for investigation and diagnosis of
PT helicase-implicated diseases such as Werner's syndrome.
XX
PS Claim 1; Page 35-42; 67pp; Japanese.
XX
CC The present sequence represents a protein having helicase activity
CC encoded by the human gene RecQ4. The gene has significant homology to the
CC Escherichia coli helicase gene (RecQ). Host cells transformed with
CC vectors comprising the RecQ4 gene are used for the recombinant expression
CC of the protein. The gene may be used for the study and diagnosis of
CC disorders in which helicase activity is involved, such as Werner's and
CC Bloom's syndromes in which mutations in the helicase gene are implicated
XX
SQ Sequence 1208 AA;

Query Match 100.0%; Score 6421; DB 2; Length 1208;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

181 ORLSLDPGWLORCHSEVDFLAGPKACRPDLGSGSESLTLPGESAVLGPAGSGGSPRAS 240
181 ORLSLDPGWLORCHSEVDFLAGPKACRPDLGSGSESLTLPGESAVLGPAGSGGSPRAS 240
241 AFOEVSIRVSGPOSSSGGKRRNMWEPWESPAQVOQESSQAGPSEBAGAVAVEDEPPG 300
241 AFOEVSIRVSGPOSSSGGKRRNMWEPWESPAQVOQESSQAGPSEBAGAVAVEDEPPG 300
301 EPVQAQPPQPCSSSPNPRYHGLSPSSQARAGABETAALHFPRLARDKRGYTVLNMKQ 360
301 EPVQAQPPQPCSSSPNPRYHGLSPSSQARAGABETAALHFPRLARDKRGYTVLNMKQ 360
301 EPVQAQPPQPCSSSPNPRYHGLSPSSQARAGABETAALHFPRLARDKRGYTVLNMKQ 360
361 KHYRGRALRSLRLRKQAKQKMRKKGCFGGGATVTTKISCFPLNEQFDHMAACCPSPA 420
361 KHYRGRALRSLRLRKQAKQKMRKKGCFGGGATVTTKISCFPLNEQFDHMAACCPSPA 420
361 KHYRGRALRSLRLRKQAKQKMRKKGCFGGGATVTTKISCFPLNEQFDHMAACCPSPA 420
421 SEEDTDVAGPEPLVPSPOPEVEPSLDPTVLPLVSLGSGQLAETPAEVPQALBQLGHQA 480
421 SEEDTDVAGPEPLVPSPOPEVEPSLDPTVLPLVSLGSGQLAETPAEVPQALBQLGHQA 480
481 PRPOGERAVMRLSGISTLVLPTGAGKSLCYQLPALIYSRRSPCLTLVSPPLSLMDQ 540
481 PRPOGERAVMRLSGISTLVLPTGAGKSLCYQLPALIYSRRSPCLTLVSPPLSLMDQ 540
541 VSGLPCLKAACIHSGMTKQRESVLOKIRAAOVHMLTPEALVAGAGLPPAQLPPVA 600
541 VSGLPCLKAACIHSGMTKQRESVLOKIRAAOVHMLTPEALVAGAGLPPAQLPPVA 600
601 PACIDEAHCLSGMSHNFRCYLAVCKVLRBMGVHCFGLTATATRTASDVACHAVAE 660
601 PACIDEAHCLSGMSHNFRCYLAVCKVLRBMGVHCFGLTATATRTASDVACHAVAE 660
661 EPDLHGPAVPVTNHLSTVSMRDTDOALLTLQGRFONDIIYYCNRREDTERIALLL 720
661 EPDLHGPAVPVTNHLSTVSMRDTDOALLTLQGRFONDIIYYCNRREDTERIALLL 720
721 RTCLHAAPVPGSGGAPKRTABAYHAGCSERRRVRQAFQOLRVVAVTAAGMGDR 780
721 RTCLHAAPVPGSGGAPKRTABAYHAGCSERRRVRQAFQOLRVVAVTAAGMGDR 780
781 PDVAVLHLGLPSPESYVQAVGRAGDGPACHLFLQPOGEDRLRLRHVHADSTDFL 840
781 PDVAVLHLGLPSPESYVQAVGRAGDGPACHLFLQPOGEDRLRLRHVHADSTDFL 840
841 AVKRLVQRPVPACTCTCTRPSPBQEGAVGERPVKTPPOABEOLSHQAFGPRRVCMGH 900
841 AVKRLVQRPVPACTCTCTRPSPBQEGAVGERPVKTPPOABEOLSHQAFGPRRVCMGH 900
901 ERALPIQLTVQALMPEBAIETLLCYELHPHHMLLATTYTCRLNCPGSPQOLQALA 960
901 ERALPIQLTVQALMPEBAIETLLCYELHPHHMLLATTYTCRLNCPGSPQOLQALA 960
961 HRCBPPLAVCLAQOLPEBPQSSSVFEDMVKLVDSMGWELASVRRALCOLQMDHEPRITGV 1020
961 HRCBPPLAVCLAQOLPEBPQSSSVFEDMVKLVDSMGWELASVRRALCOLQMDHEPRITGV 1020
1021 RRGTVLVESEBELAFHLSPGDLTAEEKQICDPLYGAVQABERQALRLRRTQAFHSV 1080
1021 RRGTVLVESEBELAFHLSPGDLTAEEKQICDPLYGAVQABERQALRLRRTQAFHSV 1080
1081 AFPSGCPCLGQODERSTRKDLGRVFESESGQPGMEBAQGEPOQALQWEDDVR 1140
1081 AFPSGCPCLGQODERSTRKDLGRVFESESGQPGMEBAQGEPOQALQWEDDVR 1140
1141 CDIROPFLSLRBEKFFSRAVARIFHGIGSPCYAQVYQODRRFMRKYLHLSFHALVGLAT 1200
1141 CDIROPFLSLRBEKFFSRAVARIFHGIGSPCYAQVYQODRRFMRKYLHLSFHALVGLAT 1200
1201 BELIQVAR 1208
1201 BELIQVAR 1208

RESULT 6
ADM90956
ID ADM90956 standard; protein; 1306 AA.
XX
AC ADM90956;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human pharmaceutically useful protein SeqID 349.
XX
KW human; cancer; haematopoiesis; thrombosis; anaemia;
KW cardiovascular disorder; ischemic heart disease;
KW acute myocardial infarction; respiratory disease; asthma; pneumonia;
KW cystic fibrosis; chronic renal failure; glomerulopathy;
KW gastrointestinal disorder; peptic ulcer; cirrhosis; immune disorder;
KW HIV infection; systemic lupus erythematosus; endocrine system;
KW diabetes mellitus; epilepsy; Alzheimer's disease;
KW amyotrophic lateral sclerosis; skin disorder; psoriasis; bacterial;
KW fungal; parasitic; viral infection; cytotoxic; anticoagulant;
KW thrombolytic; antihaemic; cardiant; vasostrictic; antiasthmatic;
KW antiinflammatory; nephrotropic; antiulcer; hepatotropic;
KW immunosuppressive; antiallergic; dermatological; antirheumatic;
KW antidiabetic; anticonvulsant; neuroprotective; nootropic;
KW antiparasitic; antibacterial; fungicide; antiparasitic; virucidal;
KW gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2004020595-A2.
XX
PD 11-MAR-2004.
XX
PF 28-AUG-2003; 2003WO-US027107.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410951P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411111P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA (RIKE-) RIKEN INST PHYSICAL & CHEM RES.
PA (DNAF-) DNAFORM KK.
XX
PI Williams LT, Chu K, Lee E, Heeitr K;
XX
DR WPI; 2004-257410/24.
DR N-PSDB; ADM90747, ADM91165.
XX
PT New human polynucleotides and polypeptides, useful for diagnosing,
PT preventing and treating proliferative disorders, immune disorders,
PT cardiovascular disorders, or bacterial, fungal, parasitic and viral
PT diseases.
XX
PS Claim 1; SEQ ID NO 349; 254pp; English.
XX
CC This invention relates to novel isolated human polynucleotides and the

CC encoded proteins thereof. Specifically, it refers to proteases, kinases,
 CC phosphatases, secreted and transmembrane proteins, as well as the derived
 CC peptide fragments, which can be used to develop antibodies and screen for
 CC small molecule agonists and antagonists that can modulate their
 CC activities. The present invention describes polypeptides,
 CC polynucleotides, vectors and host cells useful for diagnosing, preventing
 CC and treating proliferative disorders, e.g. cancer, disorders of
 CC haemopoiesis such as thrombosis and anaemia, cardiovascular disorders,
 CC e.g. ischaemic heart disease and acute myocardial infarction, respiratory
 CC diseases, e.g. asthma, pneumonia or cystic fibrosis, disorders of the
 CC kidney and urinary tract, e.g. chronic renal failure and glomerulopathy,
 CC gastrointestinal disorders, e.g. peptic ulcer or cirrhosis, immune
 CC disorders, e.g. HIV infection and systemic lupus erythematosus, disorders
 CC of the endocrine system, e.g. diabetes mellitus, central nervous system
 CC disorders, e.g. epilepsy, Alzheimer's disease or amyotrophic lateral
 CC sclerosis, skin disorders, e.g. psoriasis, as well as bacterial, fungal,
 CC parasitic and viral diseases. Accordingly, they exhibit many various
 CC activities including cytostatic, anticoagulant, thrombolytic,
 CC antinaeemic, cardiant, vasotropic, antiaesthetic, antiinflammatory,
 CC nephrotropic, anticancer, hepatotropic, immunosuppressive, antidiabetic,
 CC dermatological, antineuritic, antiarthritic, antidiabetic, antiallergic,
 CC anticonvulsant, neuroprotective, nootropic, antipsoriatic, antibacterial,
 CC fungicide, antiparasitic and virucidal, such that these polynucleotides
 CC can be used for gene therapy purposes and the development of appropriate
 CC vaccines. This polypeptide is a human protein of the invention.

XX Sequence 1306 AA;

Query Match 94.7%; Score 6085.5; DB 8; Length 1306;
 Best Local Similarity 92.6%; Pred. No. 0;
 Matches 1167; Conservative 2; Mismatches 26; Indels 65; Gaps 6;

QY 1 MERLNVRRLQWEEAFRRGRGRPSODVNAABETRALYREYTLKRTTGGAGGR 60
 DB 60 MERLNVRRLQWEEAFRRGRGRPSODVNAABETRALYREYTLKRTTGGAGGR 119
 QY 61 SSSSLPAAEEAEPFCMGPMLNRAATKSPPTPGSRQGVDPYQRLKAMIKGLQAG 120
 DB 120 SSSSLPAAEEAEPFCMGPMLNRAATKSPPTPGSRQGVDPYQRLKAMIKGLQAG 179
 QY 121 PALGRRPWPLGRASSASTPKEPGTGPVPSFAEKVSDPEPOLPEPPRGRLOHLQASIS 180
 DB 180 PALGRRPWPLGRASSASTPKEPGTGPVPSFAEKVSDPEPOLPEPPRGRLOHLQASIS 239
 QY 181 QRLGSLDPEMLQCHSEVDFPLGAPKACPDLDSESSQLIPESAVILGPGASQGPAS 240
 DB 240 QRLGSLDPEMLQCHSEVDFPLGAPKACPDLDSESSQLIPESAVILGPGASQGPAS 299
 QY 241 AFOEVSIRVGSPPSSSGGKKRMNEPWPSPAOVQOESSQAPPSGAGAVAEEDPPG 300
 DB 300 AFOEVSIRVGSPPSSSGGKKRMNEPWPSPAOVQOESSQAPPSGAGAVAEEDPPG 359
 QY 301 BPVOAOPPOPCSSPNRYHGLSPSSQARAAGKAGTAPLHPRLARHDRGNVRLNMQ 360
 DB 360 BPVOAOPPOPCSSPNRYHGLSPSSQARAAGKAGTAPLHPRLARHDRGNVRLNMQ 419
 QY 361 KHYVRRALRSRLRQAWQKRRKKGCGCGGATYTTTGESFTLNEQPDHMAAOCPPRA 420
 DB 420 KHYVRRALRSRLRQAWQKRRKKGCGCGGATYTTTGESFTLNEQPDHMAAOCPPRA 479
 QY 421 SEEDTAVGPEPLVPSPQVPEVPSLDPTLPLYSIGSPQGLAETPAEVAFOALEQGHQ 480
 DB 480 SEEDTAVGPEPLVPSPQVPEVPSLDPTLPLYSIGSPQGLAETPAEVAFOALEQGHQ 539
 QY 481 FRPGQERAVMRILSIGISTLLVLTGAGKSLCYOLPALLYRRRSPCLTLVVSPLSLMDQ 540
 DB 540 FRPGQERAVMRILSIGISTLLVLTGAGKSLCYOLPALLYRRRSPCLTLVVSPLSLMDQ 599
 QY 541 VSGLPDLCTKACIHSQMTKQRESVLOKRAOVHVMILPEALYVAGGILPPAAOLPPVA 600
 DB 600 VSGLPDLCTKACIHSQMTKQRESVLOKRAOVHVMILPEALYVAGGILPPAAOLPPVA 659
 QY 601 FACIDEAHCLSQWSHNPFCYLKCVLRERMVHCFGLTATATRTASDVAAQHLAVAE 660

DB 660 FACIDEAHCLSQWSHNPFCYLKCVLRERMVHCFGLTATATRTASDVAAQHLAVAE 719
 QY 661 EPDLHGPAVPPTNLHLSVSMRDPTDQALLTLQGRFONDSIIYCNRRDTERIALLL 720
 DB 720 EPDLHGPAVPPTNLHLSVSMRDPTDQALLTLQGRFONDSIIYCNRRDTERIALLL 779
 QY 721 RTCLHAANVPSSGGAPKTTAEAVHAGCSRRRRVQAFMGOQLR----- 766
 DB 760 RTCLHAANVPSSGGAPKTTAEAVHAGCSRRRRVQAFMGOQLRGGGGLWDGAP 839
 QY 767 -VVATVAFGKGLDREDVRAVLHLGLPSFESYQAVRAGRDGPACHLFLDP----- 820
 DB 840 ARCAAGCAAGAA---PKLRE-LRAGRGP-----GRAPLAPPPTLPYLEPPVNH 886
 QY 821 -----QGEPLRELRHVHSDTDFLAVKRLVQVFPACTCTCRPSBOGAVGGERP 873
 DB 887 TLMKVALQGBDLRELRHVHSDTDFLAVKRLVQVFPACTCTCRPSBOGAVGGERP 946
 QY 874 VPKYPPQEAQOLSHQADGPPRVCMEGRALPIQLTVQALDMPER----- 918
 DB 947 VPKYPPQEAQOLSHQADGPPRVCMEGRALPIQLTVQALDMPEREGEVSHRGVEGLS 1006
 QY 919 -----AIETVLLCTLELPHHMLLATTYTHGCLNCPGPAQIOALAHRCPPILAV 968
 DB 1007 PRPLSPALPPIETLLCYLELPHHMLLATTYTHGCLNCPGPAQIOALAHRCPPILAV 1066
 QY 969 CLAQOLPEBDPGQSSSVFEDWKLVDNMGWELASVRLACOLQWHEPRTGVRGTVLV 1028
 DB 1067 CLAQOLPEBDPGQSSSVFEDWKLVDNMGWELASVRLACOLQWHEPRTGVRGTVLV 1126
 QY 1029 EPESLAFHLRSPGDLTAAEKDOI CDPLVGRVQARERQALRLRTFTQAFHSAVPSGCP 1088
 DB 1127 EPESLAFHLRSPGDLTAAEKDOI CDPLVGRVQARERQALRLRTFTQAFHSAVPSGCP 1186
 QY 1089 LEQODEERSTRKDLIGRYFEEBEGQEPGMEDAQEPGQARLODMEDQVCDIRQFLS 1148
 DB 1187 LEQODEERSTRKDLIGRYFEEBEGQEPGMEDAQEPGQARLODMEDQVCDIRQFLS 1246
 QY 1149 LRPBEKFSRAVARLFGHIGSPCYPAQVYGQDRFWRKYLHLSFALVGLATEELLOVAR 1208
 DB 1247 LRPBEKFSRAVARLFGHIGSPCYPAQVYGQDRFWRKYLHLSFALVGLATEELLOVAR 1306

RESULT 7
 ADD01202
 ID ADD01202 standard; protein, 1144 AA.
 XX
 AC ADD01202;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human nucleic acid-associated protein NAAP-40 SEQ ID NO:40.
 XX
 KW human; nucleic acid-associated protein; NAAP; cytosolic;
 KW antiarteriosclerotic; anti-HIV; anti-allergic; cerebroprotective;
 KW antiparainflamman; anticonvulsant; nootropic; neuroprotective;
 KW antiinflammatory; ophthalmological; thyromimetic; antiarthritic;
 KW hepatotropic; antibacterial; virucide; protozoacide; antiparasitic;
 KW fungicide; gene therapy; cell proliferative disease; cancer;
 KW atherosclerosis; hepatitis; neurological disorder; Parkinson's disease;
 KW Alzheimer's disease; stroke; epilepsy; developmental disorder;
 KW renal tubular acidosis; anaemia; glaucoma; hypochyroidism;
 KW autoimmune disorder; inflammatory disorder; AIDS; allergy;
 KW atopic dermatitis; arthritis; infection.
 XX
 OS Homo sapiens.
 XX
 PN MO2003054219-A2.
 XX
 PD 03-JUL-2003.
 XX
 PF 18-DEC-2002; 2002MO-US041115.

XX 19-DEC-2001; 2001US-034306AP.
PR 11-JAN-2002; 2002US-0347633P.
PR 25-JAN-2002; 2002US-0351749P.
PR 22-FEB-2002; 2002US-0359498P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Azimzai Y, Baughn MR, Becha SD, Borowsky ML, Burford N,
PI Elliott VS, Emetling BW, Forsythe IU, Gotvad AE, Griffin JA,
PI Kadle AE, Knaie R, Lai PG, Lee EA, Lee SY, Li JX, Marquis JP,
PI Rankumar J, Richardson TW, Sprague WW, Swarnakar A, Tang YT,
PI Chawla NK, Warren BA, Yue H;
XX
XX MPI; 2003-559157/52.
DR N-PSDB; ADD01239.
XX
XX New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
PT infections.
XX
XX Claim 1; SEQ ID NO 40; 405bp; English.
XX
XX The present invention describes human nucleic acid-associated proteins
CC designated NAAP-1 to NAAP-57. The human NAAPs have cytosolic,
CC antiarteriosclerotic, anti-HIV, anti-allergic, cerebroprotective,
CC antiparkinsonian, anticoagulant, nootropic, neuroprotective,
CC antiinflammatory, ophthalmological, thymometric, antiarthritic,
CC hepatotropic, antibacterial, virucide, protozoacide, antiparasitic and
CC fungicide activities, and can be used in gene therapy. The NAAP protein
CC and polynucleotide sequences can be used in diagnosing, treating and
CC preventing diseases or conditions associated with the decreased
CC expression or overexpression of NAAP, such as cell proliferative diseases
CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
CC hypothyroidism), autoimmune/inflammatory disorders (AIDS, allergies,
CC atopic dermatitis, arthritis) and infections (e.g. bacterial, viral,
CC parasitic, protozoal, fungal). The present sequence represents human NAAP
CC -40, from the present invention.
XX
XX
XX Sequence 1144 AA;
SQ
Query Match 94.2%; Score 6050; DB 7; Length 1144;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1144; Conservative 0; Mismatches 0; Indels 64; Gaps 1;
QY 1 MERLDVVERLQAWERARFRQGRSPSODVYEAAPETTRALYREYRTLKRTTGAQGGGR 60
DB 1 MERLDVVERLQAWERARFRQGRSPSODVYEAAPETTRALYREYRTLKRTTGAQGGGR 60
QY 61 SSESILPAAAEBAPEPRCWPGLNRAATKSPQTPGRSRQSGVVDYQGLKANKLKTLOAG 120
DB 61 SSESILPAAAEBAPEPRCWPGLNRAATKSPQTPGRSRQSGVVDYQGLKANKLKTLOAG 120
QY 121 PALGRRRPMPLGRASSTKSTPKPRTGPPVPSFAEKVSDPPQLPEQPFRPGLQHLQASLS 180
DB 121 PALGRRRPMPLGRASSTKSTPKPRTGPPVPSFAEKVSDPPQLPEQPFRPGLQHLQASLS 180
QY 121 PALGRRRPMPLGRASSTKSTPKPRTGPPVPSFAEKVSDPPQLPEQPFRPGLQHLQASLS 180
DB 121 PALGRRRPMPLGRASSTKSTPKPRTGPPVPSFAEKVSDPPQLPEQPFRPGLQHLQASLS 180
QY 181 QRLGSLDPGWLQRCHEVVDLFGAPKACRPDLGSEBSQLLPGESAVLGPAGSGSGPEAS 240
DB 181 QRLGSLDPGWLQRCHEVVDLFGAPKACRPDLGSEBSQLLPGESAVLGPAGSGSGPEAS 240
QY 241 AFOEVSIRVSGSPQSSSGEGRKRNNEEPWESPAQVOQSSQAGPPSEGAIVAAVEDPPG 300
DB 241 AFOEVSIRVSGSPQSSSGEGRKRNNEEPWESPAQVOQSSQAGPPSEGAIVAAVEDPPG 300
QY 301 EPPVQAQPPQPCSSSNRPYHGLSPSSQARAKAGETAPLHFPRLARHDRGNVYRLNMKQ 360
DB 301 EPPVQAQPPQPCSSSNRPYHGLSPSSQARAKAGETAPLHFPRLARHDRGNVYRLNMKQ 360
QY 361 KHVYRGALRSRLLRKQAMKQKWRKKGCFCGGGATVTTKSCFLNBEFDHMAAQCPRPA 420

DB 361 KHVYRGALRSRLLRKQAMKQKWRKKGCFCGGGATVTTKSCFLNBEFDHMAAQCPRPA 420
QY 421 SEEDTDVAVGPEPLVPSPQVPEVPSLDPTVLPLYSIGPSGGLAETPAVEFOALBOLGHOA 480
DB 421 SEEDTDVAVGPEPLVPSPQVPEVPSLDPTVLPLYSIGPSGGLAETPAVEFOALBOLGHOA 480
QY 481 FRPQGERAVNRILSGISTLLVLPAGKSLCYQLPALLYSRRPCLTVVSPLLIMDDQ 540
DB 481 FRPQGERAVNRILSGISTLLVLPAGKSLCYQLPALLYSRRPCLTVVSPLLIMDDQ 540
QY 541 VSGLPCCAKACIHSGMTKQKQRESVLQKIRAOQVHMLPEALVGAAGLPPAQLPEVA 600
DB 541 VSGLPCCAKACIHSGMTKQKQRESVLQKIRAOQVHMLPEALVGAAGLPPAQLPEVA 600
QY 601 FACIDEAHCLSQMSHNRPCYLAVCYLRRBMGVHCEGLTATATRRRTASDVAAHVAE 660
DB 601 FACIDEAHCLSQMSHNRPCYLAVCYLRRBMGVHCEGLTATATRRRTASDVAAHVAE 660
QY 661 EPDHGRPAVPVTNLHLSVSMRDPTDQALLTLQGRKRONLDSIIYYCNREDTERIALL 720
DB 661 EPDHGRPAVPVTNLHLSVSMRDPTDQALLTLQGRKRONLDSIIYYCNREDTERIALL 720
QY 721 RTCLHAAMVPSGGRAPPTTAAVHAGMCSRERRVQRAFMQGLRVVATVAFGMGLDR 780
DB 721 RTCLHAAMVPSGGRAPPTTAAVHAGMCSRERRVQRAFMQGLRVVATVAFGMGLDR 780
QY 781 PDVRAVHLGLPPSPESVYQAVGRAGRDGPANGLFLQPGEDLRELRHHVADSTDFL 840
DB 781 PDVRAVHLGLPPSPESVYQAVGRAGRDGPANGLFLQPGEDLRELRHHVADSTDFL 840
QY 841 AVKRLVQVRVPACTCTCTRPSPSEBGAVGGERPVPKTPQGEALDSHOAAGPRVCKGH 900
DB 841 AVKRLVQVRVPACTCTCTRPSPSEBGAVGGERPVPKTPQGEALDSHOAAGPRVCKGH 900
QY 901 ERALPIQLTVQALMPBEALITTLCTYELRPHHMLLELATYTHCRINCPGPAQOLALA 960
DB 901 ERALPIQLTVQALMPBEALITTLCTYELRPHHMLLELATYTHCRINCPGPAQOLALA 960
QY 961 HRCBPPLAVCLAQQLPEDPGQSSSVFEDMYKLVDMSGMELASVRRALCOLQMDHEPRTGV 1020
DB 961 HRCBPPLAVCLAQQLPEDPGQSSSVFEDMYKLVDMSGMELASVRRALCOLQMDHEPRTGV 1020
QY 1021 RRGTVGVVRSBELAFHLRSRQDLTAERKQDCPLTGRVQAREQALRLRRTQAFHSV 1080
DB 1021 RRGTVGVVRSBELAFHLRSRQDLTAERKQDCPLTGRVQAREQALRLRRTQAFHSV 1080
QY 1077 CDIROPFLSRBEKFSRAVARIFHGIGSPCYPAQVYGODRRFRKTYLHLSFHALVGLAT 1136
DB 1077 CDIROPFLSRBEKFSRAVARIFHGIGSPCYPAQVYGODRRFRKTYLHLSFHALVGLAT 1136
QY 1201 EELLQVAR 1208
DB 1201 EELLQVAR 1208
QY 1137 EELLQVAR 1144
DB 1137 EELLQVAR 1144
RESULT 8
ID ABM84821 standard; protein; 1142 AA.
XX ABM84821;
AC
XX 18-NOV-2004 (first entry)
DT
XX Human diagnostic and therapeutic pprotein SEQ ID NO:5070.
DE
XX gene therapy; human diagnostic and therapeutic polynucleotide; dthp.
KW
XX Homo sapiens.
OS

XX MO2004023973-A2.
 XX 25-MAR-2004.
 XX 12-SEP-2003; 2003WO-US028227.
 XX 12-SEP-2002; 2002US-04102596.
 XX 12-SEP-2002; 2002US-04102606.
 XX (INCY-) INCYTE CORP.
 XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,
 XX Hartsbome TA, Sutchorolski MT, Altus CM, Pits SJ, Rider LV,
 XX Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP,
 XX Steven KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerlein EH,
 XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL,
 XX Lagace RE, Spiro PA, Stewart EA, Wingrove Y, Vilt UA, Klinton ES,
 XX Xu Y, Kwong M, Policky CJ, Hurwitz BL, Ma Y, Jackson JL, Gietzen D,
 XX Patry S, Shi X, Suarez CJ,
 XX WPI; 2004-329368/30.
 XX N-PSDB; ACN43473.
 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 XX in diagnosing a condition, disease or disorder associated with human
 XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 XX in gene mapping.
 XX Claim 27, Page; 190pp; English.
 XX The invention relates to novel diagnostic and therapeutic polynucleotides
 XX selected from one of the 2722 sequences defined in the specification. A
 XX polynucleotide of the invention may have a use in gene therapy. The human
 XX diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be
 XX used to diagnose a particular condition, disease or disorder associated
 XX with human molecules, e.g. cell proliferative disorders,
 XX autoimmune/inflammatory disorder, developmental disorders, endocrine
 XX disorder, neurological disorders, gastrointestinal disorders, or
 XX infections caused by virus, bacteria, fungi or parasite. The dthp
 XX molecules may also be used in genetic mapping, in identifying individuals
 XX from minute biological samples, in detecting single nucleotide
 XX polymorphisms, as molecular weight markers, and for somatic or germline
 XX gene therapy. The present sequence represents a dthp protein of the
 XX invention. Note: The sequence data for this patent is not represented in
 XX the printed specification, but was obtained in electronic format directly
 XX from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX Sequence 1142 AA;
 XX Query Match 93.9%; Score 6029; DB 8; Length 1142;
 XX Best Local Similarity 94.5%; Pred. No. 0;
 XX Matches 1141; Conservative 0; Mismatches 1; Indels 66; Gaps 1;
 QY 1 MERLRVRETLQWERAFRRGRSPQDVEAARETRALVREYTLKRTTQAGGGR 60
 DB 1 MERLRVRETLQWERAFRRGRSPQDVEAARETRALVREYTLKRTTQAGGGR 60
 QY 61 SSESPLAAAEAEPRMGPHILNRAATKSPQPTPGSRQGVDPYQRLKANTKGTLOAG 120
 DB 61 SSESPLAAAEAEPRMGPHILNRAATKSPQPTPGSRQGVDPYQRLKANTKGTLOAG 120
 QY 121 PALGRREPMPLGRASSASTATKPGCTGVPVFAKRVNSRPOLPEPORRRLQHLQASLS 180
 DB 121 PALGRREPMPLGRASSASTATKPGCTGVPVFAKRVNSRPOLPEPORRRLQHLQASLS 180
 QY 181 QRLGSLDPCWLRCHSEVPDFLGAAPKACRPDLGSEBSQLIPGSAVLAGGASQGEPEAS 240
 DB 181 QRLGSLDPCWLRCHSEVPDFLGAAPKACRPDLGSEBSQLIPGSAVLAGGASQGEPEAS 240
 QY 241 AFQEVSTIRVQSSPGSGGKRRWNEPWPSPPAQVOQESSQAGPPSGAGAVAVEEDPPG 300
 DB 241 AFQEVSTIRVQSSPGSGGKRRWNEPWPSPPAQVOQESSQAGPPSGAGAVAVEEDPPG 300

QY 301 EPVQAQPPQCSSPSNPRYHGLSPSSQARAKAGTAPLHIFPRLARRDRGNVYLANKKO 360
 DB 301 EPVQAQPPQCSSPSNPRYHGLSPSSQARAKAGTAPLHIFPRLARRDRGNVYLANKKO 360
 QY 361 KHYVGRALRSRLRKQAMKQKMRKKGCFGGCGATTVTTKSSCFINQEPDHAACCPRA 420
 DB 361 KHYVGRALRSRLRKQAMKQKMRKKGCFGGCGATTVTTKSSCFINQEPDHAACCPRA 420
 QY 421 SEEDTDAVGPPIVSPQVPEVPSLDPTVLPIVSLGSGGQALAPATFQALBOLGHOA 480
 DB 421 SEEDTDAVGPPIVSPQVPEVPSLDPTVLPIVSLGSGGQALAPATFQALBOLGHOA 480
 QY 481 FRPQGERAVMBILSGISTLLVLTGAGKSLCYOLPALIYSRRSPCLTIVSPLSLMDQ 540
 DB 481 FRPQGERAVMBILSGISTLLVLTGAGKSLCYOLPALIYSRRSPCLTIVSPLSLMDQ 540
 QY 541 VSGLPPLCKAACIHSQMTKRORESYLOKIRAAQVHVMLTPEALVAGAGLPPAOLPVYA 600
 DB 541 VSGLPPLCKAACIHSQMTKRORESYLOKIRAAQVHVMLTPEALVAGAGLPPAOLPVYA 600
 QY 601 FACIDBAHCLSQWSNFRPCYLKRVCKVREKGVHCFGLTATVARTASDVQHLAVAE 660
 DB 601 FACIDBAHCLSQWSNFRPCYLKRVCKVREKGVHCFGLTATVARTASDVQHLAVAE 660
 QY 661 EPDLGPAVPPTNLHLVSMDTDQALLTLLQGRFQNDIITTCNRREDTERIAALL 720
 DB 661 EPDLGPAVPPTNLHLVSMDTDQALLTLLQGRFQNDIITTCNRREDTERIAALL 720
 QY 721 RTCLHAANVPGSGGAPKTTAAVAHYAGCSRRRRRVORAFMOGOLRVVATVAFGNGLDR 780
 DB 721 RTCLHAANVPGSGGAPKTTAAVAHYAGCSRRRRRVORAFMOGOLRVVATVAFGNGLDR 780
 QY 781 PIVRAVLHLGLPSPRESYVQAVGRAGRDQPAHCHLFLQPOGEDLRELPRHYHADTDL 840
 DB 781 PIVRAVLHLGLPSPRESYVQAVGRAGRDQPAHCHLFLQPOGEDLRELPRHYHADTDL 840
 QY 841 AVKRLVQVRFPACTCTCTPRPSEOGAVGGERPVKYPQAEOLSHQAPPRVCMGH 900
 DB 841 AVKRLVQVRFPACTCTCTPRPSEOGAVGGERPVKYPQAEOLSHQAPPRVCMGH 900
 QY 901 BRALPQLTVQALDMPBEAIEITLLCYLBNPHHMLBLATYTHGCLNCGSPAPLOALA 960
 DB 901 BRALPQLTVQALDMPBEAIEITLLCYLBNPHHMLBLATYTHGCLNCGSPAPLOALA 960
 QY 961 HRCPLAVLTAOLPEDPCQSSSVFPMVKLVDSMGWLAIVRRALCOLQMDHDEPRTGV 1020
 DB 961 HRCPLAVLTAOLPEDPCQSSSVFPMVKLVDSMGWLAIVRRALCOLQMDHDEPRTGV 1020
 QY 1021 RRGTGVLVRESLAFLRSPGLTAEKQICDPLVGRVQAREBROLARLRTFOAFHSV 1080
 DB 1021 RRGTGVLVRESLAFLRSPGLTAEKQICDPLVGRVQAREBROLARLRTFOAFHSV 1080
 QY 1081 AFPSGCPCLQODEERSTRKDLKGRYFEEBSQBEFGMEDAQGPFGQARLQDWEQYR 1140
 DB 1081 AFPSGCPCLQODEERSTRKDLKGRYFEEBSQBEFGMEDAQGPFGQARLQDWEQYR 1140
 QY 1141 CDIRQTLSTRPEKSSRAVAARI FHGIGSPCTPAQVYGGDRFRMKYLLSHFALVGLAT 1200
 DB 1141 CDIRQTLSTRPEKSSRAVAARI FHGIGSPCTPAQVYGGDRFRMKYLLSHFALVGLAT 1200
 QY 1201 BELLQAVR 1208
 DB 1135 BELLQAVR 1142
 RESULT 9
 ID AAB20995
 AC AAB20995 standard; protein; 756 AA.
 XX AAB20995;
 DT 11-DEC-2000 (first entry)

XX DE Human RecQ4 helicase mutant, mut-2.
 XX KW RecQ4 helicase; human; Rothmund-Thomson syndrome; chromosome 8q24.3;
 KW poikiloderma congenitale; autosomal recessive; skin disorder;
 KW dermatology; antibody; prenatal diagnosis; gene therapy; mutant; mutain.
 OS Homo sapiens.
 XX MO200043522-A1.
 XX PD 27-JUL-2000.
 XX PF 19-JAN-2000; 2000WO-JP000233.
 XX PR 19-JAN-1999; 99JP-00011218.
 XX PA (AGEN-) AGENE RES INST CO LTD.
 PI Kitao S, Shimamoto A, Furuchi Y;
 XX WPI; 2000-524241/47.
 DR N-PSDB; AAA72365.
 XX PS Example 3; Page; 115pp; Japanese.
 XX CC The present sequence represents a mutant human RecQ4 helicase, mut-2. The
 CC cDNA encoding sequence contains a C to T substitution relative to the
 CC cDNA encoding the wild-type RecQ4 helicase (AAA72321), which changes a
 CC Gln codon (CAG) to a stop codon (TGA), causing premature truncation of
 CC the encoded protein. The invention relates to the genomic DNA sequence of
 CC human RecQ4 helicase (AAA72320). Mutations in this gene, located on
 CC chromosome 8q24.3, are the cause of Rothmund-Thomson syndrome (also known
 CC as poikiloderma congenitale), an autosomal recessive skin disorder
 CC principally occurring in females and often accompanied by juvenile
 CC cataracts, saddle nose, congenital bone defects, hypogonadism and
 CC disturbances in the growth of hair, nails and teeth. The invention also
 CC relates to vectors and host cells comprising the human RecQ4 helicase
 CC genomic sequence. It additionally encompasses use of the RecQ4 helicase
 CC protein as a therapeutic and anti-RecQ4 antibodies as diagnostic agents.
 CC The RecQ4 helicase gene and its products, and anti-RecQ4 helicase
 CC antibodies are useful in the diagnosis, especially prenatal diagnosis,
 CC and treatment of Rothmund-Thomson syndrome. The genomic sequence may
 CC especially be used in gene therapy for this condition. Note: The present
 CC sequence is not shown in the specification, but is derived from the wild-
 CC type human RecQ4 helicase shown on pages 83-92
 XX SO Sequence 756 AA;
 Query Match 62.4%; Score 4006; DB 3; Length 756;
 Best Local Similarity 100.0%; Pred. No. 2.6e-284;
 Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 AFQEVSRVSPQSSSGGKRRNNEEPWESPAYQVQSSQAQPPSBGAGVAAVEDPFG 300
 QY 301 EPVQAQPPQCSSPNPRYHGLSPSSQAQAGKAGTAPLHFPPLARHGRGNVYRLNMKQ 360
 DB 301 EPVQAQPPQCSSPNPRYHGLSPSSQAQAGKAGTAPLHFPPLARHGRGNVYRLNMKQ 360
 QY 361 KHYVGRALRSRLRKQAMQKMRKKGCCFGGGAATYTTKSCSTLNBFQDMAAQCPRPA 420
 DB 361 KHYVGRALRSRLRKQAMQKMRKKGCCFGGGAATYTTKSCSTLNBFQDMAAQCPRPA 420
 QY 421 SEEDTDAVGPPEPLVPSPPQVEVPSLDPVTYVPLVSLGPGSQLAETPAVFOALBQLGHQA 480
 DB 421 SEEDTDAVGPPEPLVPSPPQVEVPSLDPVTYVPLVSLGPGSQLAETPAVFOALBQLGHQA 480
 QY 481 FRPQGERAVMRILSGISTLVLPTGAGKSLCYQDPAALYSRRSPCLTLVSPLLSLMDQ 540
 DB 481 FRPQGERAVMRILSGISTLVLPTGAGKSLCYQDPAALYSRRSPCLTLVSPLLSLMDQ 540
 QY 541 VSGLPPLCKAACHSGMTRKQRESVLQKIRAAQVHVMLTREBALVGAGLPPAAQLPPVA 600
 DB 541 VSGLPPLCKAACHSGMTRKQRESVLQKIRAAQVHVMLTREBALVGAGLPPAAQLPPVA 600
 QY 601 PACIDEAHCLSQMSHNPFCYLRYCKVLRERMGVCFGLTATATRTTASDVAQHLAVAE 660
 DB 601 PACIDEAHCLSQMSHNPFCYLRYCKVLRERMGVCFGLTATATRTTASDVAQHLAVAE 660
 QY 661 EPDLHGAPVPTNLHLVSMDRDTQDLTLTLOGKRFQNLDSIIYYCNRREDTERIALTL 720
 DB 661 EPDLHGAPVPTNLHLVSMDRDTQDLTLTLOGKRFQNLDSIIYYCNRREDTERIALTL 720
 QY 721 RTCLHAAMVPSGGRAPTTAEATHAGMCSRRRRV 756
 DB 721 RTCLHAAMVPSGGRAPTTAEATHAGMCSRRRRV 756
 RESULT 10
 ID ADW44021
 ID ADW44021 standard; protein; 1216 AA.
 AC ADW44021;
 XX 24-MAR-2005 (first entry)
 DT Mouse RECQL4 protein.
 DE
 XX RECQL4; helicase; protein deactivation; growth disorder;
 KW Rothmund-Thomson syndrome; skin disorder; bone disease;
 KW musculoskeletal disease; enzyme.
 XX Mus musculus.
 OS
 XX WO200501085-A1.
 PN 06-JAN-2005.
 PD 25-JUN-2004; 2004WO-JP009380.
 XX 27-JUN-2003; 2003JP-00185409.
 PR
 XX (AGEN) NAT INST RADIOLOGICAL SCT.
 PA
 XX Abe M;
 PI WPI; 2005-081681/09.
 DR N-PSDB; ADW44020.
 XX Novel non-human mammal, preferably rodent such as mouse exhibiting
 PT characteristics of Rothmund-Thomson syndrome, bearing RECQL4-gene lacking
 PT mouse having mutation in RECQL4-gene, useful as model of human Rothmund-
 PT Thomson syndrome.
 XX Disclosure; SEQ ID NO 2; 69pp; Japanese.

XX This invention describes a novel knock-out mouse which lacks the RECQ4 gene or has a mutation in one of exons 13 to 22 of the RECQ4 gene resulting in a loss of helicase activity. The mouse described in the invention exhibits the characteristics of Rothmund-Thomson syndrome and is suitable as a model of human Rothmund-Thomson syndrome. The mouse exhibits growth delay, skin abnormality and bone-formation defects similar to human Rothmund-Thomson syndrome. This sequence represents a murine RECQ4 which is missing in mice exhibiting Rothmund-Thomson syndrome.

XX Sequence 1216 AA:

61.4%: Score 3944; Db 9; Length 1216;

Query Match Best Local Similarity 63.9%: Pred. No. 1.7e-279; Matches 795; Conservative 110; Mismatches 267; Indels 72; Gaps 13;

```

QY 1 MERLDVBERLQAWERARFRQRGRRRSODVEAAEETRALYREYRLKRTTQAGGGLR 60
DB 1 MERLATVABRLQEWERAFAFLHGRPAKGVDAEAEETRALYREYRLKQAVRQADNHR 60
QY 61 SSE-SLPPAAEBAEPKPCGPHNAAATKSQPTGRSGQSVPDYQGLRANLKTIOA 119
DB 61 VLEOSLAABEAQBPSCMGPHLSRAATQNTQSMKQSLSSVDQGRKLANLKNITQT 120
QY 120 GPALGRRRWPILGRASSKASTPKPQGTGPVSPFAKVSSEPPQLPRPPRPGRLQIASL 179
DB 121 GPTGSRKQLOKRLSTVPAPRPSPKTESPCPDADALPRVPRPRRLQLOOLRSSL 180
QY 180 SQRLGSLDPGWLQRCHEVPPDPLGAPKACRPDLGSEBSQLIPGESAVLGPAGS----Q 235
DB 181 SRRLSLDPGWLERCHNRVSDLEVPAGCGLDLSAEBSQPMGSKRVNIADPDIQSEVSQ 240
QY 236 GPEASAFQEVSTIRVSGPOSSSGEKRWNEEPWESPAVOQESSQAPSPSGAGAVAE 295
DB 241 SPEAIAQOPAVQLSOSPSPKSKGRKRKNEK-GDFPQ-DQPSGAGPLSGAATATYHG 298
QY 296 BDPPEBPVQAQBPPOPCSSPNRHYGLSPSOARAGKAGTAPLHPPLARHNRGVYR 355
DB 299 QDPPEBPVQVNPQPCNS-----SNQARTEKAGITHLASPPRPSLDKGNVIR 347
QY 356 LMMKQHYVGRALSRLLRKQAMQKWKKECEGCGGATVTTESCLNQPDMMAQ 415
DB 348 LMMKRRFVRVAGANGRLRKQVMQKWKQAAGSGSPRATDQTCRCGQFGHMASQ 407
QY 416 CPRPA-----SEEDTAVGP-----EPLVSPQVPEVPSL 446
DB 408 CSQPGPTLVQEGEDRDQPISTLEEVAAQRTGTASCHNSGEGTQPAAPELQ-VPHCP- 465
QY 447 DPTVPLVLSGPGSLAETPAEVPQALBOLGHOAPRPGQERAVMRLSGISTLLVLPSTA 506
DB 466 --PMSPLVPPGPGVQALTPAEVFOALERLGYRAFPQGERALMRLSGISTLLVLPSTA 523
QY 507 GKSICVLPALVYSRSPCLTVVSPFLSLAMDQVSGLPPLCAACIHSGMTKQRESVL 566
DB 524 GKSICVLPALVYAGRSPCLTVVSPFLSLAMDQVSDPLSCAKACIHSGMTKQRESVL 583
QY 567 QKIRPAQVHVMLTPEALVGA-----GLPRAAQLPVPAFACIDEAHCLISQMSHNPFCY 621
DB 584 KKVRAAQVHVMLTPEALVGA-----GLPRAAQLPVPAFACIDEAHCLISQMSHNPFCY 643
QY 622 LAVCVLBERMGVHGLTATARTASDVNOHLVAEERPLHGAPVPTLHLISVSD 681
DB 644 LAVCVLBERMGVHGLTATARTASDVNOHLVAEERPLHGAPVPTLHLISVSD 703
QY 682 RPTDQALTLTLQKRFONLSIIYCNRRDETERIALRLTCLHAAMVPSGGRAPKTTA 741
DB 704 RPSDQALTLTLQKDRRTLDSTVIYCTBRIRIQNGMLATLRLTCLSMWGDSRPGCGEALTA 763
QY 742 EAYHAGMGRERRRRVQARFMQGLRVVATVAFGMGLDPRVRAVTHLGLPSPFSEYVA 801
DB 764 EAYHAGMSSQERRRRVQARFMGLRMVAVTVAFGMGLDPRVRAVTHLGLPSPFSEYVA 823

```

```

QY 802 VGRAGRDGPAHCHLFLQPGSEDLRELRHHVHADSTDFLAVKRLVQVFPACTCTTRPP 861
DB 824 IGRAGRGKPAHCHLFWHPQGEDLMELRRHHAADSTDFLAVKRLVQVFPACTCTCTCT 878
QY 862 SEQBGAVGGRPPVPKYRPPQEAQSL-HQA PGPRVCMGERALPIQLTYQALDMPBEAI 920
DB 879 -----QRPVSKSSPEEVKESGGQTYPVLGQAQCLGHERALPVOSTVQALMTBEAI 929
QY 921 ETLICYLELPHHMLLELTATYTHCRINCPGPAQLOALAHRCPLAVCLAAQOLPEDPGQ 980
DB 930 ETLICYLELPHHMLLELTATYTHCRINCPGPAQLOALAHRCPLAVCLAAQOLPEDPGQ 989
QY 981 GSSSEEDMVKLVDSMGWELASVRAALCOLQMDHEPRTGVRRGTGLVVERSELAFLHRSR 1040
DB 990 GRSLLEGGVVELADSMGKTLASVQAHLQKMDPEPKGAAGQGVLVKFSLELAFHLHSR 1049
QY 1041 GDLTAEKQDCDLYNGRVQAREQALALRLRTQAFHSVAFPSGCGCLTEQDDEBSRL 1100
DB 1050 GDLTAEKQDCDLYNGRVQAREQALALRLRTQAFHSVAFPSGCGCLTEQDDEBSRL 1109
QY 1101 KDLGRYFEEBEGQEPGAMEDAQPPEPQARLOPMEQVRCDIQFLSLRPEKFSRAV 1160
DB 1110 KTLVSYTFEEBE-EEBETMTDTQPKFGQTLQDMEQDIRDVQSLSTRBEHFSGRAV 1168
QY 1161 ARIFHGIGSPCYPAQVYGODRRFWKRYLHLSFHALVGLATEELL 1204
DB 1169 ARIFHGIASPCYPAQVYGLDRFRFKYLLHDFHALMLHATEELL 1212

```

RESULT 11

AA020994 ID AA020994 standard; protein; 554 AA.

AA020994; 11-DEC-2000 (first entry)

Human RecQ4 helicase mutant, mut-1.

RecQ4 helicase; human; Rothmund-Thomson syndrome; chromosome 8q24.3; polylodermis congenitalis; autosomal recessive; skin disorder; dermatology; antibody; prenatal diagnosis; gene therapy; mutant; mutcin.

Homo sapiens.

WO200043522-A1.

27-JUL-2000.

19-JAN-2000; 2000WO-JP000233.

19-JAN-1999; 99JP-00011218.

(AGEN-) AGENE RES INST CO LTD.

Kitao S, Shimamoto A, Furuichi Y;

WPI; 2000-524241/47.

N-PSDB; AAA72364.

RecQ4 helicase gene, gene products and antibody, used in the diagnosis and treatment of Rothmund-Thomson syndrome, e.g. by gene therapy.

Example 3; Page; 115pp; Japanese.

The present sequence represents a mutant human RecQ4 helicase, mut-1. This cDNA encoding sequence contains a 7 bp deletion relative to the cDNA encoding the wild-type RecQ4 helicase (AA072321), which alters the reading frame and causes premature truncation of the encoded protein. The invention relates to the genomic DNA sequence of human RecQ4 helicase (AA072320). Mutations in this gene, located on chromosome 8q24.3, are the cause of Rothmund-Thomson syndrome (also known as polylodermis congenitalis), an autosomal recessive skin disorder principally occurring

in females and often accompanied by juvenile cataracts, saddle nose, congenital bone defects, hypogonadism and disturbances in the growth of hair, nails and teeth. The invention also relates to vectors and host cells comprising the human RecQ4 helicase genomic sequence. It additionally encompasses use of the RecQ4 helicase protein as a therapeutic and anti-RecQ4 antibodies as diagnostic agents. The RecQ4 helicase gene and its products, and anti-RecQ4 helicase antibodies are useful in the diagnosis, especially prenatal diagnosis, and treatment of Rothmund-Thomson syndrome. The genomic sequence may especially be used in gene therapy for this condition. Note: The present sequence is not shown in the specification, but is derived from the wild-type human RecQ4 helicase shown on pages 83-92

Sequence 554 AA;

Query Match 45.5%; Score 2920; DB 3; Length 554;
Best Local Similarity 100.0%; Pred. No. 6.5e-205;
Matches 550; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MERLRDVERIQAWERARFRGRGRPPSODDVEAAPEERLALYREYRTLKRTTGGAGGGLR 60
1 MERLRDVERIQAWERARFRGRGRPPSODDVEAAPEERLALYREYRTLKRTTGGAGGGLR 60
61 SSESIPAAAEAPRRCWCPHLNRAATKSPQTPGRSRQSVDPYQRLKANTLKTGLAQ 120
61 SSESIPAAAEAPRRCWCPHLNRAATKSPQTPGRSRQSVDPYQRLKANTLKTGLAQ 120
121 PALGRPPPLGRASSKASTPKPRTGTPVPSFAEKVSDPEPQLPEPQPRRLQHLQASLS 180
121 PALGRPPPLGRASSKASTPKPRTGTPVPSFAEKVSDPEPQLPEPQPRRLQHLQASLS 180
181 QRLSLDGMVLRCHSEVPDLGAPKACRPLDGSSEQLLPGESAVLGPAGSGGPBAS 240
181 QRLSLDGMVLRCHSEVPDLGAPKACRPLDGSSEQLLPGESAVLGPAGSGGPBAS 240
241 AFOEVSIRVSGPOSSSGEGRKRWNEEPWESPAVOQSSQAGPSEAGAVAVEDEPG 300
241 AFOEVSIRVSGPOSSSGEGRKRWNEEPWESPAVOQSSQAGPSEAGAVAVEDEPG 300
301 EPPVQAOPPOPCSSPSNPRYHGLSPSSQARAGAKGTADEPLAHPRLARRDRGNVYRLNMKQ 360
301 EPPVQAOPPOPCSSPSNPRYHGLSPSSQARAGAKGTADEPLAHPRLARRDRGNVYRLNMKQ 360
361 KHYVRGALRRLRKQWKQKMKKRCGREGGATVTTTKSCPLNEPDMAAQCPPEA 420
361 KHYVRGALRRLRKQWKQKMKKRCGREGGATVTTTKSCPLNEPDMAAQCPPEA 420
421 SEEDTDVAGPEPLVSPQVPEVPSLDPTVLPLYSIGSGQLAETPAVFOALBQLGHQA 480
421 SEEDTDVAGPEPLVSPQVPEVPSLDPTVLPLYSIGSGQLAETPAVFOALBQLGHQA 480
481 FRPQOEBAVMRIISGISTLLVLTGAGKSLCYQALPALLYSRRSPCLTLVSPPLSLMDMQ 540
481 FRPQOEBAVMRIISGISTLLVLTGAGKSLCYQALPALLYSRRSPCLTLVSPPLSLMDMQ 540
541 VSGLPPLCKA 550
541 VSGLPPLCKA 550

RESULT 12

ABB63317 ID ABB63317 standard; protein; 1579 AA.

AC ABB63317;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 16743.

KM Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li FWD, Myers EW,

XX WPI; 2001-656860/75.

XX N-PSDB; ABL07420.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

XX Disclosure; SEQ ID NO 16743; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signaling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1579 AA;

Query Match 26.8%; Score 1723.5; DB 4; Length 1579;
Best Local Similarity 28.9%; Pred. No. 1.3e-116;
Matches 467; Conservative 217; Mismatches 459; Indels 471; Gaps 45;

10 RLQAWERAPRGRGRPPSODDVEAAPEERLALYREYRTLKRT----- 51
16 RLVKWERDQFKKGNRVPSKYDIRDSQERDSYKMYKLTSPLEETLNDVLSGDYDL 75
52 -TGQAGC-----GLASS--ESLP----- 66
76 EMSQASDFGVSMLDQVSLNEGRLPLDIALVGPQSSGNLEIPQVSGSFNTLDLPP 135
67 -----AAAEAP-EPRCWCPHLNRAATKSPQ-----TPGRSRQG 100
136 RQVLTNLVNDENHVIKRFVAVELPINQNAWGLNVAKPPAPQVPEASKAGHGKQ 194
101 SVRPYQRLKANTLKTGL-QAGPALGR--TWPLGRASSKASTPKPRTGTPVPSFAEKVSD 157
195 --PFAAGSLKPSLSAKLFQSSRGPAKKNPKKPLSRCVSSSSSTTSLGSLVPLDHHLELDF 252
158 EPPQLPEPQPRPGRLO-----HLQASLSQRLGSLDPPGLQR-----CHSEVP- 199
253 ETLIRKAQRYKKEQCALANNPMLASHESKESIKT---LYDDGRLKRYNTENTLDEVLPF 309
200 ----DFVGAPE-----ACRPDLGSESSQLLIPESAVL-----GP 230
310 AEAANNNGTSKKTNFGILANLDSLKTPTVKEKVLQAKPPQMAIIQELQTDNMSMNQKP 369
231 GAGSQPELASAFQEVSRVSGPQSS-----SGEGRKRWNE----- 267
370 DHLNHPASSQSVAKR-NKPPSEGETSDSDSVVAESEBQEPQETRLQSKRRKIVS 428
268 -----PWESPAVOQES---SQAGP----- 284
429 TASQKVEVAAPVELPNKVEPTEFAQENDFSADEQDQATYVPENKKQKAKRQAKG 488

```

QY 285 -----PSEGAAGAAVEEDPGEVQAOPOP----- 310
DB 489 QUTTKPKAEKPKTEKKAAYAKKPKAKPKRNSKKAIAVAPADPEDEKQPLNPEL 548
QY 311 -----CSPSPNPRYH-----GLSPSQARAKAEGTAP 338
DB 549 KYVLALEAGDITSVPRIINVQDLEADATAGRYIRTPAAGPNPGLSEGSNIRVDEKKAAR 608
QY 339 LHIFFPLAHD-RGNVYRLNKKYKHYRG-RALRSLLAKOMKOKR-----KKG 387
DB 609 KKLERRIAAGKLNENVTJNIOCKKVRKKSYSFKYKQQRHKKRVAALSGPDMWG 668
QY 388 ECFGGGATVTTTKESCFLEQFDMHAAOC-----PRP----- 419
DB 669 GCDGG-----VLT-----CFGGGVHFPQOCKYKGDLSLLPLSAQLEEDPSPPTLAEGEMA 721
QY 420 -----ASEEDTAVGPEPLVSPQP----- 439
DB 722 SGGAVAHNRNISRLLPQAAANAAILQGBELNESEEDQESGDEEYQOHDPNMSDEMDV 781
QY 440 -----VPE-----VPELDP-----VL 451
DB 782 DFEALDAVEASISQVSOEKAASPIKTYVGAKIPEEFLKQAGLDTTASSNRSQHGK 841
QY 452 PLVSLPQSGLAETPAVEFQALBOLGHOAFRPGQERAVNRILSGISTLLVLPAGAKSLC 511
DB 842 PLVLDLPDGSVQDTTEBEVLALMFQHTNPKKQDRAIKRTLSGLSVLTLGSGSKSLC 901
QY 512 YQPLALLYSRSPCLTLVSPPLSLMDQVSGLPCLCAKACHSGMTRKRESVLQKIRA 571
DB 902 YQPLALLYSRSPCLTLVSPPLSLMDQVSGLPCLCAKACHSGMTRKRESVLQKIRA 571
QY 572 AQVHTMLTPBALY-----GAGCLPPAOLPVPAFACIDBAACLSQMSNRPCTLRVC 625
DB 962 GERDILVSPBEAVABERATFGAI--LRQLPRIAFACIDBAACVQMSNRPCTLRVC 1019
QY 626 KVLREHGVCFGLTATATRTASDVAOHLAAV--BPDLGAPAPNTLHLVSDMDRT 684
DB 1020 KVLKRLGVTVGLTATATLTPRVSLINHLGSDERGIISDLPDLNVLVSVDENR 1079
QY 685 DQALLTLLOKRFQNDLSIIYCNRRREDTERIALRTCLHAAMVPSGGRAKPT-----T 740
DB 1080 DALLTLQINSERFEPQCSIIITCRDEGERIAGFRTQODRREFPODTKKRKNVWQ 1139
QY 741 AEAHAGMCRERRRQVORAMOGOLRVYVATVAFGMLDPRDVRALVHLGPPSPFSYVQ 800
DB 1140 AEAHAGMCRERRRQVORAMOGOLRVYVATVAFGMLDPRDVRALVHLGPPSPFSYVQ 800
QY 801 AVBRAGDQOPACHLPLPOGDBLRBRHVADSTDFLAVKRLVORVPACTCTCTRP 860
DB 1200 EIGRAGDGLPSRCHLPLDAGKQDSELRHVYNSIDRVHTRKLDKTVPCSC----- 1254
QY 861 PSEOGAVGGERVUPKYPQEBARQLSHQAAPGR-RVCMGHERALPQLTVQALMDPEEA 919
DB 1255 --DKMSKRTALP-----LEBDGRVHMCPPHREIGFSYKTVEMLDIPAEN 1300
QY 920 IETLLCYLELHPHMELELATTTTHCLNCPGPAQOLALNRCPLAVCLAQOLPEPD- 978
DB 1301 ISTLLCTMEALDPRWCISVLSSAYMAVVISYGGPKYLKHAKECPPLAMAIALQIRKTF 1360
QY 979 GQGSASSVEPMVKLVDSMGWELASVVRALQOLMDHEPRTGVRRGTGVLEFSELAHRL 1038
DB 1361 KEBDSNITFESVTDIAGIGMSGVKQLQDLER--VKVMGYPKRSITVSPDLGRIRK 1418
QY 1039 SPQDLTAEBKQOICDPLVGRVQARERQALRLRTTFAFHSVAPSPGCLBQO-DEERS 1097
DB 1419 VPGDFTSESIDNALDILYTRSVKQERTQLQIYVAHGLAAVASSGQCCNADFPQDRG 1478
QY 1098 TRKMDLGRYEEBEGQ-----BRGMEBAQGEPEQANLQDMEDVQDCDIRFSLRP 1151
DB 1479 EQLFAIVRNPANDYPQDLELELEPSNVPD-----ENIIDVHALINMYP 1523
QY 1152 EEKPSRAVAIRIFGIGSPCYPAQVYGDRFRWRKYLHLSPHALVGLATELLQ 1205

```

```

DB 1524 DNTFTGRNARIIFHIGIMSPNYPVAVIWR-CRFRVAHVCDPFRILHLNMAIRK 1576
RESULT 13
AAU34821
ID AAU34821 standard; protein; 610 AA.
XX
AC AAU34821;
XX
DT 14-FEB-2002 (first entry)
XX
DE E. coli cellular proliferation protein #402.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX
OS Escherichia coli.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
XX
PR 23-MAY-2000; 2000US-0206848P.
XX
PR 26-MAY-2000; 2000US-0207727P.
XX
PR 23-OCT-2000; 2000US-0242578P.
XX
PR 27-NOV-2000; 2000US-0253625P.
XX
PR 22-DEC-2000; 2000US-0257931P.
XX
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
XX DR N-PSDB; AAS52680.
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 10414; 511bp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence represents an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 610 AA:
XX
Query Match 9.8%; Score 631.5; DB 4; Length 610;
Best Local Similarity 37.6%; Pred. No. 3, 6e-37;
Matches 164; Conservative 63; Mismatches 134; Indels 75; Gaps 15;
QY 474 EQLGHAFRPGQERAVNRILSGISTLLVLPAGAKSLCYQLPALLYSRSPCLTLVSP 533

```


AC ABU28286;
 XX 19-JUN-2003 (first entry)
 DT Protein encoded by prokaryotic essential gene #13813.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Enterobacter cloacae.
 OS WO200277183-A2.
 XX 03-OCT-2002.
 PD 21-MAR-2002; 2002MO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zykkind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI, 2003-029926/02.
 DR N-PSDB; ACA32156.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 56210; 1766bp; English.
 PS The invention relates to an isolated nucleic acid comprising any one of
 XX the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 609 AA:
 Query Match 9.6%; Score 619; DB 6; Length 609;
 Best Local Similarity 35.4%; Pred. No. 2, 9e-36;
 Matches 169; Conservative 73; Mismatches 152; Indels 84; Gaps 18;

QY 474 EOLGHOAFRPGOERAVNKRLLSGISTLLVLPPTGAKSKSLCYOPLALLYSRRSPCLTLVSP 533
 DB 19 ETFGQGRPGOERTLIETVLEGRDCLVMPPTGGKSLCYVPALVLNG----LTVVSP 74
 QY 534 LSLMDQVSG-LPPLCKRACHSGMTRKQRESYVQKTRPAQVHYMLTPALVAGGLPP 592
 DB 75 ISLMKDQVDQLANGVAAACINSTQTRQOQEVNAGCRVQIRLMYLAPEELMDNFIDH 134
 QY 593 AAQLPVAFACIDEAHCLSGSHNFRPCYLRVCYLREMGVHGLGLTATATRTASDV 652
 DB 135 LAHNPVLLA-VDEAHCTSGMHDPRPYALGO-LKRFPELPPMALITATADTTLDI 192
 QY 653 AQHLAAVEEPDLHGPAPVPTMHLV-SMDRDTQALLTLQGRFPQNLDSI----- 703
 DB 193 VRLIG-----LNDP-----YIQVSSFDPNIRYMLM-----EKFKPLDQLLRVVGQR 235
 QY 704 ----IYCNREDTERIALIRTCLHAAMVPGSGRAKTTAEVYHAGMCSRRRRVQRA 759
 DB 236 GKSGLIYCNRSRAKVEDTAARLQN-----RGFSAAAYHAGLEHHIRADVOEK 281
 QY 760 FMQGLRVVAVTVARGMGLDRPDVRAVLHLGLPSFESYVQAVGRAGRDGPACHLFL- 818
 DB 282 FORDDLQIVAVTVARGMGLDRPDVRAVLHLGLPSFESYVQAVGRAGRDGPACHLFL- 341
 QY 819 -----OPGEDLRELRRH---VHADSTDFLAVRLV-----QVFPACTC 855
 DB 342 PADMAKLRCLBEKRGQ-LODIERHKLNMGAFAEAQTCRLVILNFGESRGRCPCNC 400
 QY 856 -TCTRPSEQAGAVGERPVVK-YPPQAEQLSH-----QAAPEFRVCMGHERALPT 906
 DB 401 DICLPPKQYDGLMDARALSTIYVNRQFGMGVVEVLRGANNRIRIDMGHDK-LPV 457

Search completed: December 27, 2005, 21:48:06
 Job time : 224.8 secs

THIS PAGE BLANK (USPTO)